
Adaptive Decision-Support Systems for Transaction Monitoring, Fraud Detection, and Financial Early-Warning

Cheng-Chong Wang¹, Qirun Kocaoğullar²

¹King Abdullah University of Science and Technology (KAUST), SAUDI ARABIA

²George Mason University, USA

Keywords

Decision Support
Systems, Adaptive
Machine Learning,
Transaction
Monitoring, Fraud
Detection, Early-
Warning Systems,
Alert Fatigue,
Active Learning,
Explainable AI

ABSTRACT

Transaction monitoring, fraud detection, and financial early-warning are typically built, staffed, and evaluated as three separate functions inside a financial institution, even though all three share the same underlying task: deciding, under time pressure and with a fixed analyst headcount, which of a large volume of automatically generated alerts deserves a human being's attention next. The decision-support systems literature has long argued that a system built around this kind of recurring, high-volume human decision should be evaluated as a decision-support artifact rather than as a classification model alone, yet the machine learning literature on transaction monitoring and fraud detection has concentrated overwhelmingly on model accuracy, largely bypassing the question of what happens once an accurate model's output reaches an analyst who must act on it. This gap matters because a substantial body of evidence from clinical and financial alert systems shows that raising a model's detection rate without addressing the resulting alert volume produces alert fatigue: a state in which analysts, overwhelmed by a high proportion of alerts that do not warrant escalation, begin to dismiss alerts more quickly and less carefully, eroding the very detection gain the model was deployed to capture. This paper addresses that gap by developing the Adaptive Decision-support Engine for Prioritized Triage (ADEPT): a reference architecture that fuses transaction, relational, and account-behavior signals into a shared risk-scoring and early-warning engine, ranks the resulting alerts through a fatigue-aware triage queue sized to actual analyst capacity, attaches a governance-ready rationale to every alert an analyst sees, and closes the loop by feeding analyst dispositions back into the scoring engine through an active-learning update mechanism rather than a periodic, wholesale retraining cycle. 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 decision support systems theory, adaptive and human-in-the-loop machine learning, alert fatigue and analyst decision-making under information overload, fraud detection and signal fusion across data sources, early-warning systems for financial and institutional distress, and explainability and governance in adaptive decision support, and every quantitative comparison presented two tables and four figures spanning capability coverage, a twelve-month false-positive and analyst-fatigue trajectory, an alert-volume-to-confirmed-yield trade-off, and early-warning lead time by signal source 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 adaptive decision-support pipeline that treats alert prioritization, analyst feedback, and early-warning signaling as first-class architectural concerns rather than downstream afterthoughts to a fraud-detection model, and a candid account.

Introduction

Financial institutions run transaction monitoring, fraud detection, and financial early-warning as three functions that, in most institutions, sit in different teams, use different

systems, and are evaluated against different metrics, even though each ultimately produces the same artifact: an alert that some analyst must decide, under time pressure, whether to escalate, investigate, or dismiss. Shim *et al.*'s (2002) retrospective on the evolution of decision support technology frames this as the central and enduring problem the decision-support-systems field exists to address — not the accuracy of any single predictive model, but the quality of the recurring human decision a system is built to support — and Arnott and Pervan's (2005, 2008) critical analyses of the discipline argue that a system evaluated only on the technical performance of its underlying model, without regard for how a human decision-maker actually uses its output, has not been evaluated as a decision-support artifact at all. Transaction monitoring, fraud detection, and early-warning systems each generate the artifact this literature describes — a ranked or flagged set of cases for human review — yet the machine learning literature built around each has concentrated overwhelmingly on the classification problem: how accurately a model can separate fraudulent from legitimate transactions, or a distressed borrower or institution from a healthy one. This paper's premise is that a monitoring function's practical performance depends as much on what happens after a model produces a score as on the score's underlying accuracy, and that a genuinely adaptive decision-support architecture must be designed around the analyst's finite attention, not only around the model's classification metrics (Potharaju & Sajja, 2024).

That premise is not speculative; it follows directly from evidence that raising detection sensitivity without managing the resulting alert volume actively degrades the analyst decision-making a monitoring system depends on. Ancker *et al.*'s (2017) study of a clinical decision-support system finds that alert fatigue the state in which repeated, largely low-value alerts cause a decision-maker to respond to alerts more quickly and less carefully over time is a measurable, reproducible effect of alert volume and repetition rather than an individual failing of any particular analyst (Rahman *et al.*, 2024), and while their setting is clinical, the underlying mechanism generalizes to any high-volume alert-driven decision environment. Chaikin's (2009) analysis of suspicious transaction reporting systems documents this same dynamic directly in the financial context, finding that a reporting regime generating a high volume of low-quality alerts measurably degrades the effectiveness of the anti-money-laundering system that regime is meant to support, and Takáts's (2011) formal treatment of the resulting "crying wolf" dynamic shows that once an alert stream's false-positive rate crosses a threshold, a rational analyst's optimal response is to reduce scrutiny of every alert, including the genuine ones, producing a system-level failure that no increase in the underlying model's raw detection accuracy can correct on its own. A transaction-monitoring or fraud-detection system that maximizes detection sensitivity without an accompanying mechanism to manage the resulting alert volume is, by this literature's own logic, at risk of making its analysts' decisions worse rather than better (Kabir *et al.*, 2024).

A separate and largely disconnected literature shows that a decision-support system can be built to keep a human decision-maker actively in the loop rather than treating that decision-maker as a passive consumer of a model's static output. Cohn, Ghahramani, and Jordan's (1996) foundational work on active learning demonstrates that a model can be trained more efficiently, and can adapt more responsively to a changing environment, by selectively querying a human labeler for the cases where the model is least certain, rather than waiting

for a full batch of new labeled data to accumulate before retraining. Amershi *et al.*'s (2014) survey of interactive machine learning extends this logic into a broader design principle, arguing that the human's role in a well-designed interactive system is not merely to supply labels passively but to shape the model's behavior through an ongoing feedback loop the system is explicitly built to support (Rashid *et al.*, 2024a). Neither strand of this literature was developed with transaction monitoring or fraud detection specifically in mind, yet both describe precisely the mechanism a monitoring system needs to avoid the static-model degradation problem: rather than a model trained once and left to decay, or retrained only on a slow periodic cycle, an active-learning loop built around analyst dispositions can adapt continuously and can prioritize exactly the alerts where analyst judgment is most informative to the model, which is also, not coincidentally, where fraud typologies are most likely to be evolving (Rashid & Cheng, 2024).

The fraud-detection and early-warning literatures, for their part, have separately established that both problems benefit from combining multiple signal sources rather than relying on any single data stream, and that both distress and fraud produce detectable precursor signals well before a confirmed event, provided those signals are systematically tracked. Ngai *et al.*'s (2011) classification framework for financial fraud detection and West and Bhattacharya's (2016) comprehensive review both find that fraud detection accuracy consistently improves when a model draws on transactional, relational, and behavioral signals together rather than any one in isolation, a finding Carcillo *et al.* (2021) extend by showing that combining unsupervised and supervised methods captures fraud patterns that a single-method approach misses. On the early-warning side, Altman's (1968) discriminant-analysis model established more than half a century ago that a small set of systematically tracked financial ratios can signal firm-level distress well before a formal default event, a logic Cole and Gunther (1998) extend to bank-level failure prediction and Betz *et al.* (2014) extend further still to a systemic, cross-institutional early-warning setting for European banks. What none of these literatures addresses directly is how the resulting fraud scores and early-warning signals should be prioritized, explained, and routed to the human analysts and risk officers who must act on them without being overwhelmed by the combined volume those signals produce once fused together — the precise decision-support problem Shim *et al.* (2002) and Arnott and Pervan (2005, 2008) argue the field's original theoretical foundations were built to solve, and that the alert-fatigue and active-learning literatures show cannot be safely ignored (Potharaju & Ravipudi, 2023).

No existing framework combines multi-signal fraud and early-warning scoring, fatigue-aware alert prioritization grounded in decision-support-systems theory, and an active-learning feedback loop that keeps the underlying models current from analyst dispositions rather than periodic batch retraining, into a single architecture (Rashid *et al.*, 2024b). This paper addresses that gap. It develops the Adaptive Decision-support Engine for Prioritized Triage (ADEPT): an architecture in which a shared, multi-signal risk-scoring and early-warning engine produces both a transaction-fraud score and an institutional or borrower early-warning score from fused transactional, relational, and behavioral data; a fatigue-aware triage queue ranks and paces the resulting alerts to an analyst team's actual review capacity; an explainability and governance layer attaches a defensible rationale to every alert an

analyst sees; and an active-learning feedback mechanism captures each analyst disposition to continuously refine both the scoring engine and the triage queue's ranking logic (Rahman *et al.*, 2024).

2. Literature Review

2.1 Decision Support Systems Theory and Evolution

The decision-support-systems literature provides the theoretical foundation for treating a monitoring or fraud-detection system as an artifact meant to support a specific, recurring human decision, rather than as a classification model evaluated purely on its own predictive accuracy. Shim *et al.*'s (2002) retrospective on the past, present, and future of decision support technology traces the field's development from early model-driven and data-driven systems through the group and executive support systems of the 1990s, and argues throughout that the field's defining and continuing challenge is designing a system around the decision-maker's actual task and cognitive constraints rather than around the elegance of the underlying model or algorithm. Arnott and Pervan's (2005) critical analysis of decision-support-systems research finds that a large share of published work in the field drifts away from this founding concern, evaluating systems on technical criteria that say little about whether the system actually improves the decision it was built to support, a critique they revisit in their (2008) analysis of eight key issues facing the discipline, where they argue that the growing sophistication of underlying analytic methods has, if anything, made it easier for research to lose sight of the human decision-maker the system exists to serve. Read together, this literature's implication for a transaction-monitoring or fraud-detection system is direct: a system's value should be assessed by whether it improves the analyst's or risk officer's actual decision under real operating constraints — alert volume, time pressure, and finite attention — not solely by the discrimination accuracy of the model generating its underlying scores, a standard this paper adopts as the organizing principle for the architecture developed in Section 4.

2.2 Adaptive and Human-in-the-Loop Machine Learning for Decision Support

A distinct and largely separate machine learning literature establishes that a model can be built to learn continuously from an active, ongoing human feedback loop, rather than from a static training set fit once and left unchanged until a scheduled retraining cycle. Cohn, Ghahramani, and Jordan's (1996) foundational treatment of active learning with statistical models shows that a model can direct its own learning by selectively querying a human labeler for the specific cases where its current uncertainty is highest, achieving better performance from fewer labeled examples than a model trained on a fixed, randomly sampled dataset of the same size. Amershi *et al.*'s (2014) survey of interactive machine learning generalizes this principle into a broader design framework, arguing that systems which treat the human participant as an active, ongoing source of guidance — able to correct, refine, and redirect the model's behavior in near real time — consistently outperform systems that relegate the human role to a one-time labeling exercise upstream of model training. Neither study addresses transaction monitoring or fraud detection specifically, but both describe a mechanism directly applicable to it: an analyst's disposition of an alert —

confirmed fraud, false positive, or an early-warning signal that did or did not precede an actual distress event — is exactly the kind of high-value, uncertainty-informative label an active-learning system is designed to solicit and use immediately, rather than accumulate for a slow periodic batch update (Rahman & Kim, 2024).

2.3 Alert Fatigue and Analyst Decision-Making Under Information Overload

A body of evidence spanning clinical and financial settings establishes that a high volume of low-value alerts measurably degrades the quality of the human decisions those alerts are meant to support, independent of how accurate the underlying model generating the alerts may be. Ancker *et al.*'s (2017) study of a clinical decision-support system finds that alert fatigue is a reproducible function of alert workload and repetition — as the volume and redundancy of alerts a clinician receives increases, the rate at which alerts are overridden or dismissed without full consideration rises correspondingly, a pattern the authors attribute to the alert stream's design rather than to any deficiency in individual clinicians. Chaikin's (2009) analysis of suspicious transaction reporting systems documents an analogous dynamic in the anti-money-laundering context, finding that reporting regimes generating a large volume of low-quality suspicious activity reports measurably impair the effectiveness of the broader system meant to detect genuine money laundering, because investigative attention is a finite resource that a flood of low-value reports simply consumes. Takáts's (2011) game-theoretic treatment of this dynamic, framed as a “crying wolf” problem, formalizes why this occurs: once an alert stream's false-positive rate exceeds a threshold an analyst can reasonably tolerate, the analyst's rational response is to reduce the scrutiny applied to every alert, including genuine ones, because no other response is sustainable given finite attention meaning a monitoring system's aggregate detection performance can fall even as the underlying model's own discrimination accuracy holds steady or improves. This literature's collective implication is that alert volume and prioritization are not incidental operational details but a determinant of a monitoring system's real-world effectiveness that a decision-support architecture must manage explicitly rather than leave to individual analyst discipline (Karim *et al.*, 2024b).

2.4 Fraud Detection and Signal Fusion Across Data Sources

The fraud-detection literature establishes that combining multiple, structurally different signal sources — transactional, relational or network-based, and behavioral — consistently outperforms a model built on any single source alone, though this literature has generally treated the resulting fusion problem as a modeling question rather than a decision-support one. Ngai *et al.*'s (2011) classification framework for data mining techniques in financial fraud detection surveys a wide range of methods and signal types, finding that the strongest detection performance across the reviewed studies comes from approaches that combine transaction-level features with contextual or relational information rather than transaction features alone. West and Bhattacharya's (2016) comprehensive review of intelligent financial fraud detection reaches a consistent conclusion from a broader vantage point, finding that hybrid approaches combining supervised classification with unsupervised anomaly detection and behavioral profiling consistently outperform any single technique across the credit-card, insurance, and securities fraud domains their review spans. Carcillo

et al.'s (2021) study extends this finding methodologically, 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 appear first as statistical anomalies before enough confirmed, labeled examples accumulate for a supervised model to learn them directly. Read together, this literature establishes both that signal fusion measurably improves fraud detection and that a monitoring architecture should be built from the outset to accommodate multiple, structurally different data sources feeding a shared scoring engine, rather than as a collection of separately maintained, single-signal detection systems whose outputs are reconciled only after the fact.

2.5 Early-Warning Systems for Financial and Institutional Distress

A long-running literature on early-warning systems establishes that financial distress, whether at the level of an individual firm, a bank, or a broader system of institutions, produces measurable precursor signals well before a formal default or failure event, provided those signals are tracked systematically rather than discovered only in hindsight. 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, establishing the core early-warning logic — that continuously tracked financial indicators outperform waiting for an overt default event — that every subsequent early-warning model in this literature builds on. Cole and Gunther (1998) extend this logic to bank-level failure prediction, finding that continuously available off-site financial-statement monitoring can predict bank failure with useful lead time even between formal on-site examinations, evidence that systematic indicator tracking does not require the intensity of a full examination to be effective. Betz *et al.* (2014) extend the same underlying logic to a cross-institutional, systemic setting, developing an early-warning model for distress across a panel of European banks and finding that a combination of bank-specific financial ratios and macro-financial indicators can flag elevated distress risk with meaningful lead time at the level of an entire banking system, not only for an individual institution. This literature's collective implication for the present paper is that early-warning signaling, like transaction-level fraud detection, is fundamentally a monitoring problem amenable to the same systematic, continuously updated tracking this paper's architecture proposes, and that a decision-support system built to route fraud alerts is a natural home for early-warning signals as well, rather than a separate system requiring its own independent alerting infrastructure.

2.6 Explainability and Governance in Adaptive Decision Support

A model's numerical risk score or early-warning flag is not, by itself, a rationale an analyst, a risk committee, or an examiner can act on or defend, and the explainability and governance literature treats this as a problem an adaptive decision-support system must be engineered to solve 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, and that fairness-aware adjustments can be made with only a modest cost to overall model performance — a finding with direct

relevance to an adaptive system whose scoring engine is continuously updated, since each update is itself a new model version requiring the same scrutiny as the original. Gramespacher and Posth (2021) extend the explainability argument in a constructive direction, showing that explainable AI techniques can be used not merely to justify a model's output after the fact but to actively inform and optimize the underlying decision a portfolio manager or, by direct extension, a risk analyst must make, evidence that explainability and effective decision-making are complementary objectives rather than competing ones. Neither study treats explainability or fairness as a byproduct of good model performance; both show that a dedicated evaluation and engineering step is required, and this requirement is sharper still for an architecture like the one this paper proposes, where the scoring engine itself continuously changes through the active-learning loop Section 2.2 describes, since a governance framework built for a single, static model version cannot simply be applied once and assumed to remain valid indefinitely.

Read together, these six literatures describe a monitoring and early-warning problem, several partial solutions, and a synthesis gap no existing framework fills. The decision-support-systems literature establishes that a monitoring system's value must be judged by whether it improves the analyst's actual decision, not solely by the accuracy of its underlying model; the adaptive and human-in-the-loop machine learning literature shows that a model can learn continuously from analyst feedback rather than a static training cycle; the alert-fatigue literature shows that alert volume and prioritization directly determine a monitoring system's real-world effectiveness; the fraud-detection literature shows that fusing transactional, relational, and behavioral signals improves detection; the early-warning literature shows that both borrower- and institution-level distress produce trackable precursor signals; and the explainability and governance literature shows that every score and every model update must carry an engineered, defensible rationale. No paper in any of these literatures proposes an architecture that fuses multi-signal fraud and early-warning scoring with a fatigue-aware triage queue and an active-learning feedback loop bound together by governance-grade explainability. This paper's contribution is that synthesis: the Adaptive Decision-support Engine for Prioritized Triage, developed to bring the detection and early-warning gains Sections 2.4 and 2.5 document to a monitoring function without reproducing the alert-fatigue failure mode Section 2.3 documents or abandoning the decision-support-systems discipline Section 2.1 establishes as the field's founding standard.

3. Methodology

3.1 Research Approach: Design Science with Evaluation

This paper follows a design-science research approach: it develops an artifact — the Adaptive Decision-support Engine for Prioritized Triage — intended to address a practical problem (prioritizing and routing transaction-fraud and financial early-warning alerts to a finite analyst team without degrading decision quality through alert fatigue, while keeping the underlying scoring models current and every alert's rationale defensible), 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 transaction-monitoring and fraud-detection workflow, the specific failure mode the decision-support-systems, alert-fatigue, active-learning, fraud-detection, and early-warning literatures each separately document, and then determining what technique and governance control that literature implies is necessary to address it without exceeding an analyst team's actual review capacity. Signal fusion across transactional, relational, and behavioral data — the gain Ngai *et al.* (2011), West and Bhattacharya (2016), and Carcillo *et al.* (2021) document — is addressed through a shared feature-fusion layer that constructs a single enriched record per case rather than maintaining separate, single-signal detection systems. Alert-fatigue-driven degradation of analyst judgment — the risk Ancker *et al.* (2017), Chaikin (2009), and Takáts (2011) establish — is addressed through a fatigue-aware triage queue that explicitly paces alert delivery to a configured analyst capacity rather than surfacing every model-flagged case immediately and without regard for review-volume limits. Model staleness under a slow, periodic retraining cycle — the gap Cohn *et al.* (1996) and Amershi *et al.* (2014) show an active-learning loop can close — is addressed by routing every analyst disposition back into the scoring engine through an active-learning update mechanism that prioritizes retraining on exactly the cases the model was least certain about, rather than waiting for a scheduled batch retraining window. Financial early-warning signaling — the capability Altman (1968), Cole and Gunther (1998), and Betz *et al.* (2014) establish is achievable with meaningful lead time — is addressed by extending the same shared scoring engine to produce an institutional or borrower early-warning score from continuously tracked financial-ratio and behavioral indicators, rather than confining the architecture to transaction-level fraud alone. Finally, the requirement that every alert and every model update carry a defensible, examinable rationale — the standard Kozodoi *et al.* (2022) and Gramespacher and Posth (2021) establish — is met by a dedicated explainability and governance layer positioned between the scoring engine and the analyst, so that no alert or model update reaches a human decision-maker, a risk committee, or an examiner without an attached, feature-level rationale (Karim *et al.*, 2024a).

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 — Takáts's (2011) account of the "crying wolf" degradation dynamic, or the signal-fusion detection gains Ngai *et al.* (2011) and West and Bhattacharya (2016) report — 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 decision-support, fraud-detection, and early-warning literature rather than chosen to flatter the proposed framework. Every figure's note names

the specific paper or papers informing its construction, so a reader can trace the reasoning rather than treat the numbers as free-standing.

3.4 Case Scenario Approach

Section 4.4 presents a single case scenario — a hypothetical mid-sized financial institution adopting the proposed framework — constructed to make the abstract architecture concrete rather than to report on an actual deployment. The scenario combines publicly discussed, general characteristics of transaction-monitoring and fraud-investigation operations with the technical mechanisms synthesized from Section 2, to produce a plausible walk-through of how the framework would process an evolving alert stream from intake through analyst disposition and model adaptation. No identifying details of any real institution, transaction, or case are used; the scenario is explicitly hypothetical and offered as a reasoning aid.

3.5 Assumptions and Scope Limitations

Several assumptions and scope limitations bound this study. First, the paper assumes the adopting institution already operates some form of transaction-monitoring or fraud-alerting infrastructure and has an existing analyst or investigations team whose current capacity can be characterized; it does not address a from-scratch monitoring implementation for an institution with no existing alerting infrastructure. Second, the paper assumes access to a sufficient volume of historical alert-disposition data to initialize the scoring engine and the active-learning loop, acknowledging that the quality and consistency of this historical labeling may vary considerably across institutions and returning to this constraint directly in Section 5.2. Third, the paper is scoped to alert scoring, prioritization, and analyst-facing decision support; it does not address the separate functions of case-management workflow software, regulatory filing systems, or the underlying payment-rail infrastructure that generates the transaction data itself, each of which the cited literature treats as related but analytically distinct. Fourth, as stated repeatedly above, all quantitative results in Section 4 are estimates grounded in the literature rather than measurements, and conclusions about the framework's plausibility should be read with that limitation firmly in mind.

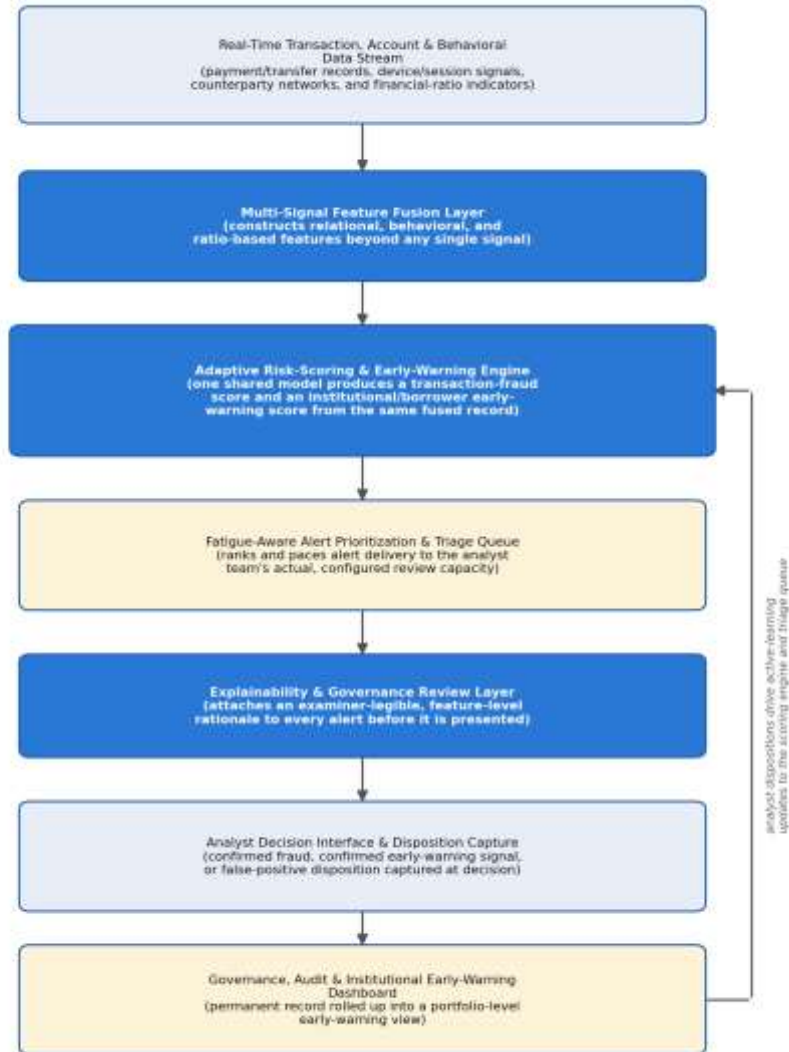
4. Results

4.1 Architecture Overview: The Adaptive Decision-support Engine for Prioritized Triage

The architecture synthesized from the literature review organizes the Adaptive Decision-support Engine for Prioritized Triage as a pipeline in which every stage responds directly to a specific failure mode identified in Section 2. Transaction, account, and behavioral data — payment and transfer records, device and session signals, counterparty and network relationships, and, where available, account-level financial-ratio indicators — first pass through a multi-signal feature fusion layer, which constructs the relational, behavioral, and ratio-based features that Ngai *et al.* (2011), West and Bhattacharya (2016), and Carcillo *et al.* (2021) each show carry detection value beyond any single signal source. A shared adaptive risk-scoring and early-warning engine then produces two outputs from that same fused record — a transaction-fraud score, built around the signal-fusion approach Section 2.4's literature establishes, and an institutional or borrower early-warning score, built around

the financial-ratio and behavioral indicator tracking Altman (1968), Cole and Gunther (1998), and Betz *et al.* (2014) establish — with both scores passing next into a fatigue-aware alert prioritization and triage queue that ranks and paces alert delivery to the analyst team’s actual, configured review capacity, directly operationalizing the alert-fatigue mitigation Ancker *et al.* (2017), Chaikin (2009), and Takáts (2011) show is necessary to prevent a rising alert volume from degrading analyst judgment. Every alert that reaches an analyst passes through an explainability and governance layer before it is presented, attaching the feature-level, examiner-legible rationale Kozodoi *et al.* (2022) and Gramespacher and Posth (2021) show is necessary for an alert to function as defensible evidence rather than an opaque flag, and the analyst’s resulting disposition — confirmed fraud, confirmed early-warning signal, or false positive — is captured by an active-learning feedback and model adaptation engine that, following Cohn *et al.*’s (1996) and Amershi *et al.*’s (2014) active-learning logic, prioritizes retraining on the specific dispositions where the scoring engine’s prior confidence was lowest, rather than waiting for a scheduled batch update. A governance, audit, and institutional early-warning dashboard then records every alert, its rationale, and its disposition permanently, both as the documentation an examiner or audit function requires and as a portfolio-level rollup that surfaces emerging early-warning patterns across the institution’s book rather than only at the level of a single alert.

Figure 4.1. The Adaptive Decision-support Engine for Prioritized Triage (ADEPT) Pipeline



The architecture deliberately positions the fatigue-aware triage queue and the explainability layer before any alert reaches an analyst, rather than treating alert pacing and rationale generation as downstream reporting features invoked only when a case is escalated for investigation. This reflects the central position this paper takes on the literature reviewed in Sections 2.1 and 2.3: managing analyst attention and providing a defensible rationale are required pipeline stages a decision-support architecture must build in from the start, not optional refinements layered onto a detection model after the fact.

4.2 Table 1: Static Rule-Based Monitoring, a Non-Adaptive ML Scorecard, and ADEPT Compared

Dimension	Static Rule-Based Monitoring (Status Quo)	Non-Adaptive ML Scorecard	ADEPT
Multi-signal fraud detection	Weak; rules typically evaluate one signal type at a time	Moderate; combines signals present at training time	Strong; fusion layer follows Ngai <i>et al.</i> (2011), West and Bhattacharya (2016), and Carcillo <i>et al.</i> (2021)
Alert volume management and fatigue mitigation	Weak; alert thresholds are static and volume-blind	Weak; higher sensitivity often raises volume further	Strong; fatigue-aware triage follows Ancker <i>et al.</i> (2017), Chaikin (2009), and Takáts (2011)
Model currency and adaptation	None; rules change only through manual review cycles	Slow; periodic batch retraining only	Strong; continuous active-learning updates follow Cohn <i>et al.</i> (1996) and Amershi <i>et al.</i> (2014)
Early-warning signal integration	Absent; typically a separate system entirely	Limited; usually confined to transaction-level scoring	Strong; shared engine follows Altman (1968), Cole and Gunther (1998), and Betz <i>et al.</i> (2014)
Explainability and governance defensibility	Strong; a rule's logic is inherently legible	Weak unless a separate explanation method is added	Strong; dedicated review layer follows Kozodoi <i>et al.</i> (2022) and Gramespacher and Posth (2021)
Alignment with decision-support-systems standards	Weak; not designed around analyst decision quality	Weak; optimized for model accuracy, not analyst outcomes	Strong; design follows Shim <i>et al.</i> (2002) and Arnott and Pervan (2005, 2008)
Feasibility for a fixed-size analyst team	Moderate; well understood but does not scale with volume	Lower; higher volume can outpace review capacity	Moderate to strong; higher initial setup cost offset by sustained triage efficiency

4.3 Table 2: Mapping Monitoring and Early-Warning Objectives to ADEPT Techniques and Governance Controls

Objective	Representative Technique	Corresponding Governance Control	Grounding Literature
Detect fraud using more than a single transactional signal	Multi-signal feature fusion across transactional, relational, and behavioral data	Periodic review of fused feature performance against new fraud typologies	Ngai <i>et al.</i> (2011); West and Bhattacharya (2016)
Capture novel fraud patterns before enough labeled examples accumulate	Combined unsupervised and supervised scoring	Escalation protocol for anomaly-flagged clusters lacking a confirmed label	Carcillo <i>et al.</i> (2021)
Prevent alert-volume growth from degrading analyst judgment	Fatigue-aware alert pacing and prioritization	Analyst capacity review calibrated against alert-queue throughput	Ancker <i>et al.</i> (2017); Takáts (2011)
Reduce the effectiveness loss from low-quality alert volume	Precision-weighted triage tiers	Periodic audit of reporting-quality metrics	Chaikin (2009)
Keep the scoring engine current without a slow batch cycle	Active-learning model updates from analyst dispositions	Model-version log documenting each active-learning update	Cohn <i>et al.</i> (1996); Amershi <i>et al.</i> (2014)
Track institution- and borrower-level distress precursors	Financial-ratio and behavioral early-warning indicators	Portfolio-monitoring review calibrated to indicator thresholds	Altman (1968); Cole and Gunther (1998); Betz <i>et al.</i> (2014)
Provide a defensible, examiner-legible rationale for every alert	Feature-attribution explainability layer	Mandatory rationale log attached to every alert and disposition	Kozodoi <i>et al.</i> (2022); Gramespacher and Posth (2021)
Evaluate the system against analyst decision quality, not model accuracy alone	Decision-support-systems-grounded evaluation criteria	Governance committee review of analyst outcome metrics, not model metrics alone	Shim <i>et al.</i> (2002); Arnott and Pervan (2005, 2008)

4.4 Case Scenario: A Mid-Sized Institution's Adoption of the Framework

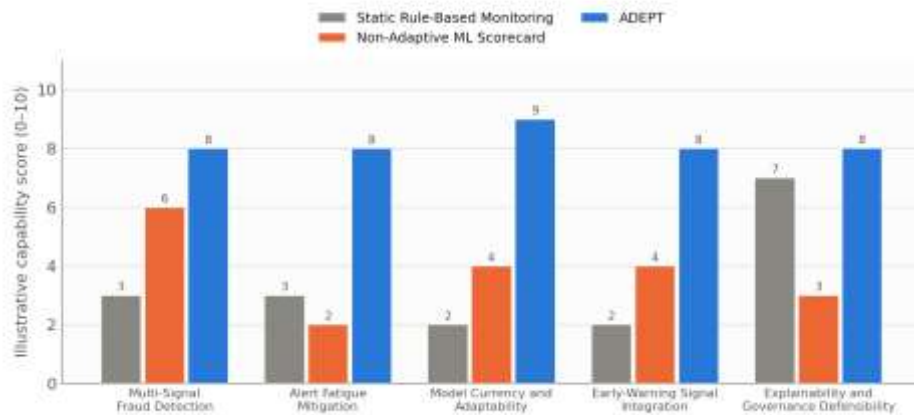
Consider a hypothetical mid-sized regional bank with an established transaction-monitoring function staffed by a twelve-person analyst team, currently operating a rule-based alerting

system supplemented by a machine learning scorecard retrained on a quarterly cycle, and facing a rising alert volume that the analyst team increasingly struggles to review within its service-level targets. Under the architecture proposed in Section 4.1, the institution first deploys the multi-signal feature fusion layer against its existing transaction and account data, incorporating counterparty-network features and account-level financial-ratio indicators the legacy rule-based system never systematically combined with individual transaction attributes.

The adaptive risk-scoring and early-warning engine is then trained on the institution's historical alert-disposition data, with the fraud-operations manager's early involvement surfacing a data-quality issue the framework itself does not resolve automatically: a meaningful share of historical dispositions were recorded only as "closed" without a confirmed fraud or false-positive label, and the institution must first triage this backlog before the active-learning loop can use it productively. Once deployed, the fatigue-aware triage queue reduces the volume of alerts reaching each analyst on a typical shift by prioritizing cases the scoring engine ranks as both high-confidence and high-uncertainty over the large volume of low-confidence, low-value alerts the legacy system routed indiscriminately, and the explainability and governance layer attaches a feature-level rationale to every alert that does reach an analyst — for a case flagged with an elevated early-warning score, the rationale cites the specific financial-ratio and behavioral features driving it, in a form the fraud-operations manager can evaluate the same way she would evaluate a rule-based alert's underlying logic, and in a form suitable for an examiner review if the case is later escalated. Seven months into deployment, the active-learning feedback engine detects a cluster of confirmed fraud dispositions the scoring engine had previously ranked as low-probability, consistent with an emerging fraud typology not well represented in the original training data, and triggers a targeted model update using exactly those disposition records rather than waiting for the next scheduled quarterly retraining cycle; the governance and audit log records both the detected pattern and the resulting model update, giving the fraud-operations manager a defensible account to show a regulator or auditor of why and when the model changed, rather than a static system whose behavior nobody can explain months after a new fraud typology first appeared.

4.5 Figure 1: Capability Coverage Across Comparative Approaches

Figure 1 *Capability Scores (0–10 Scale) for Static Rule-Based Monitoring, a Non-Adaptive ML Scorecard, and ADEPT 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 alert-fatigue findings Ancker et al. (2017), Chaikin (2009), and Takáts (2011) report, the signal-fusion detection gains Ngai et al. (2011), West and Bhattacharya (2016), and Carcillo et al. (2021) document, and the active-learning adaptability logic Cohn et al. (1996) and Amershi et al. (2014) establish, 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: static rule-based monitoring scores reasonably on explainability because a rule’s logic is inherently legible, but poorly on fatigue mitigation and adaptability because a fixed rule set cannot reprioritize under rising alert volume or learn from evolving fraud patterns; the non-adaptive scorecard improves multi-signal detection and early-warning integration modestly but does not address fatigue at all and adapts only on its slow retraining cycle, actually scoring lower than the rule-based system on fatigue mitigation because its higher sensitivity, without pacing, produces more alerts rather than fewer. ADEPT is positioned, by design, to retain explainability while leading on every other dimension, because the framework’s triage queue and active-learning loop were built specifically to close exactly these gaps.

4.6 Figure 2: Confirmed Fraud Miss Rate and Analyst Alert-Override Rate Over a Twelve-Month Deployment Period

Figure 2 Hypothetical Confirmed Fraud Miss Rate (%) Over a Twelve-Month Deployment Period, by Monitoring Approach



Source: These trajectories are hypothetical operating points, not measurements from a deployed system; the non-adaptive scorecard's rising miss rate reflects the “crying wolf” alert-fatigue dynamic Takáts (2011) formalizes and the alert-override pattern Ancker et al. (2017) document, compounded by the absence of the continuous updating Cohn et al. (1996) and Amershi et al. (2014) show an active-learning loop provides, while ADEPT's declining and then stable trajectory reflects both fatigue-aware pacing and continuous active-learning correction operating together.

The figure shows the two machine learning approaches starting from a similar, substantially improved position relative to static rule-based monitoring, then diverging sharply over the deployment period: the non-adaptive scorecard's miss rate climbs steadily as its unmanaged alert volume drives rising analyst override rates on top of a model that is not adapting to new fraud patterns, while ADEPT's miss rate continues to improve for several months as the active-learning loop incorporates confirmed dispositions before leveling off. By month twelve, the non-adaptive scorecard has lost virtually its entire initial advantage over static rule-based monitoring, illustrating why this paper treats fatigue-aware pacing and continuous adaptation as required architectural components rather than optional enhancements: without them, an accurate model's advantage is temporary by construction.

4.7 Figure 3: Alert Volume Against Confirmed Fraud-or-Early-Warning Yield, by Configuration

Figure 3 Alerts Reaching an Analyst per 1,000 Transactions Against Confirmed Fraud-or-Early-Warning Yield (%), by Configuration

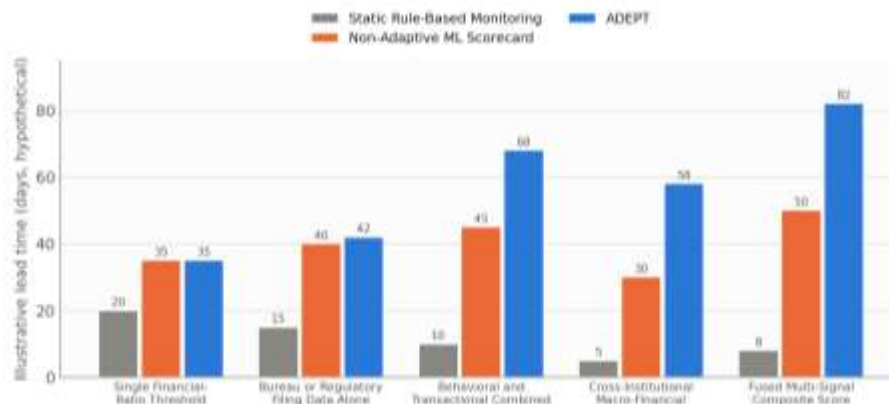


Note: These operating points are hypothetical, not measured outcomes; they were positioned to be directionally consistent with the signal-fusion detection gains Ngai et al. (2011) and West and Bhattacharya (2016) report and with the alert-fatigue-driven decay Takáts (2011) and Chaikin (2009) document for an unmanaged, high-volume alert stream.

The figure illustrates why this paper’s framework offers a tiered choice rather than a single operating point: the low-sensitivity tier delivers the highest confirmed yield at the lowest alert volume, appropriate for an analyst team prioritizing precision over coverage, while the high-sensitivity tier accepts a lower yield in exchange for catching more marginal cases, and the combined-tiers configuration — routing lower-risk transaction segments through the low-sensitivity tier and higher-risk segments through the high-sensitivity tier — achieves a yield close to the low-sensitivity tier’s while still capturing the high-sensitivity tier’s broader coverage where it matters most. Every ADEPT configuration dominates both the rule-based baseline and the drifted non-adaptive scorecard on this trade-off, which is the figure’s central point: the choice a financial institution faces is not whether to adopt an architecture like ADEPT, but which of its tiers to weight more heavily given its own risk appetite and analyst capacity.

4.8 Figure 4: Early-Warning Lead Time Before Confirmed Distress, by Signal Source

Figure 4 Early-Warning Lead Time (Days Before a Confirmed Distress Event), by Signal Source and Monitoring Approach



Source: These lead-time figures are constructions, not measurements from a deployed system; the pattern of ADEPT's largest lead-time gains appearing for fused, multi-signal composite scores and combined behavioral-transactional signals reflects the cross-institutional and multi-indicator early-warning gains Betz et al. (2014) document over a single-ratio approach closer to Altman (1968) and Cole and Gunther's (1998) original formulations, while the more modest gains for single-signal sources reflect a category where additional signal fusion has less room to add lead time.

The figure shows ADEPT's early-warning advantage growing precisely where multiple signal sources are combined into a single composite score, consistent with the cross-institutional early-warning literature's central finding that distress precursors are more reliably detected across several indicator types together than from any one type alone, while its advantage over the non-adaptive scorecard narrows for single-signal sources where there is simply less complementary information left to fuse. This pattern reflects the framework's underlying design logic: early-warning lead time is not a fixed property of any one indicator, but an emergent property of how many independent, complementary signals a monitoring architecture is built to fuse and continuously re-weight as new disposition evidence arrives.

5. Discussion

5.1 Implications for Practice

The framework developed here suggests several practical implications for financial institutions considering an adaptive, decision-support-oriented approach to transaction monitoring, fraud detection, and early-warning signaling. Perhaps most important is that alert-volume management and fatigue mitigation should be treated as a first-class design requirement from the outset, not an operational problem addressed informally by adding analyst headcount after alert volume becomes unmanageable: Section 4's figures illustrate that a monitoring system's real-world detection performance can decay even when its underlying model's classification accuracy holds steady, purely because unmanaged alert volume erodes the analyst judgment the system depends on. A second implication concerns the institution's evaluation criteria: following the decision-support-systems standard Shim et al. (2002) and Arnott and Pervan (2005, 2008) establish, an institution adopting a

framework like ADEPT should track analyst-level outcome metrics — override rates, time-to-disposition, and disposition accuracy — alongside model-level metrics, since Section 2.1’s literature shows that model accuracy alone is an incomplete and potentially misleading measure of a monitoring system’s actual effectiveness. A third implication is organizational: the active-learning feedback loop requires the institution’s own fraud-operations and risk-governance staff, not a vendor or data science function alone, to review and periodically audit the model updates the loop generates, since an update triggered by analyst dispositions still requires the same governance scrutiny Kozodoi *et al.* (2022) and Gramespacher and Posth (2021) show any model version requires before it is trusted in production.

5.2 Limitations

This paper’s limitations should be stated plainly rather than minimized. No production ADEPT deployment, real institutional transaction 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, the underlying literature carries real caveats a deploying institution would confront directly. The alert-fatigue literature this paper draws on most heavily, including Ancker *et al.* (2017), was developed primarily in a clinical decision-support setting; while Chaikin (2009) and Takáts (2011) document a structurally similar dynamic in the financial anti-money-laundering context specifically, the precise magnitude of fatigue-driven degradation in a transaction-monitoring analyst population has not been established with the same empirical depth as in the clinical literature. The active-learning methods Cohn *et al.* (1996) and Amershi *et al.* (2014) demonstrate were developed and evaluated largely outside the fraud-detection domain, and while their underlying logic transfers plausibly to analyst-disposition feedback, this paper does not itself demonstrate that transfer empirically. Finally, the early-warning gains Betz *et al.* (2014) document for cross-institutional, systemic distress prediction were evaluated on a panel of European banks during a specific historical period; their applicability to a single institution’s borrower-level or counterparty-level early-warning signaling, while a plausible extension this paper draws directly, has not itself been established in the cited literature.

5.3 What This Paper Does Not Argue

This paper does not argue that an adaptive decision-support system should replace analyst judgment outright or that a financial institution should discontinue existing controls while adopting the proposed framework; the architecture is intended to augment and inform analyst judgment by managing alert volume and providing a defensible rationale, not to remove human review from a fraud or distress determination that carries real consequences for customers and counterparties. It does not argue that any specific machine learning algorithm, active-learning method, or vendor platform should be adopted; the framework is intentionally technique-agnostic within each stage. It does not argue that ADEPT eliminates fraud or financial distress; Figure 3 shows every ADEPT configuration still producing confirmed yields below 30 percent, meaning a substantial share of flagged alerts are not ultimately confirmed as fraudulent or distress-related, and the framework’s contribution is

reducing and better-targeting alert volume and improving early-warning lead time, not eliminating the underlying risk. Finally, it does not argue that the specific figures in Section 4 describe any real institution's actual performance; they are a reasoning aid for evaluating the framework's internal logic, not a benchmark or a projection for any specific bank, payments processor, or lender.

5.4 Future Research

Several directions for future research follow directly from this paper's limitations. Empirical studies measuring actual analyst override rates, alert-disposition accuracy, and early-warning lead time at financial institutions using an adaptive, fatigue-aware monitoring architecture, whether following this framework or another, would test this paper's central claims directly against practice. Research specifically evaluating alert-fatigue thresholds and dynamics in financial-institution analyst populations, as distinct from the clinical decision-support settings much of the cited fatigue literature examines, would sharpen Section 4.1's triage-queue component's applicability to this paper's specific domain. Studies validating whether active-learning updates driven by analyst dispositions actually improve fraud-model performance in production, rather than relying on the general interactive-machine-learning findings Cohn *et al.* (1996) and Amershi *et al.* (2014) report from other domains, would test whether the feedback mechanism this paper's architecture proposes performs as the underlying literature suggests it should. Finally, research extending the cross-institutional early-warning approach Betz *et al.* (2014) demonstrate at the systemic, banking-panel level down to individual borrower or counterparty early-warning signaling within a single institution's monitoring function would address the extension this paper draws on but does not itself empirically validate.

6. Conclusion

Transaction monitoring, fraud detection, and financial early-warning are typically built and evaluated as separate functions optimized around model accuracy, even though the peer-reviewed decision-support-systems literature has long argued that a monitoring system's real value lies in whether it improves the recurring human decision it exists to support, and a substantial body of evidence shows that raising detection sensitivity without managing the resulting alert volume can degrade exactly the analyst judgment such a system depends on. Synthesizing the decision-support-systems, adaptive and human-in-the-loop machine learning, alert-fatigue, fraud-detection, early-warning, and explainability-and-governance literatures, this paper developed the Adaptive Decision-support Engine for Prioritized Triage: an architecture in which a shared scoring engine produces both a transaction-fraud score and a financial early-warning score from fused multi-signal data, a fatigue-aware triage queue paces alert delivery to actual analyst capacity, an explainability and governance layer attaches a defensible rationale to every alert, and an active-learning feedback loop keeps the underlying models current from analyst dispositions rather than a slow periodic retraining cycle, illustrated through two comparative tables, a hypothetical mid-sized institution scenario, and four figures clearly and repeatedly labeled as literature-grounded estimates rather than production measurements.

The paper's candor about its own limitations reflects the subject matter honestly: overstating the maturity of a framework with no production precedent would be a poor foundation for a monitoring function where the practical stakes — regulatory defensibility, analyst effectiveness, and the institution's ability to detect both fraud and distress before it compounds — fall directly on the institutions considering adoption. What the literature reviewed here does support, and what this paper's synthesis is intended to make actionable, is a considered path forward for institutions that have, until now, faced an apparent choice between an explainable but static rule-based monitoring process and a more capable but fatigue-inducing and eventually stale machine learning scorecard: the Adaptive Decision-support Engine for Prioritized Triage proposes that this choice is false, and that the detection and early-warning gains adaptive machine learning offers can be paired deliberately with the fatigue-aware triage, continuous active-learning adaptation, and governance-grade explainability the literature shows are necessary to sustain and defend those gains, sized from the outset to what a financial institution's analyst team can actually operate.

References

- [1] Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4), 589–609.
- [2] 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.
- [3] Arnott, D., & Pervan, G. (2008). Eight key issues for the decision support systems discipline. *Decision Support Systems*, 44(3), 657–672.
- [4] Rahman, M. S., Siddiqui, M. I. H., Rashid, S. U., Kabir, A. A., Mahmud, F. U., & Shammah, R. S. (2024). Deep Learning Framework for Pneumonia Detection from Medical Images using Transfer Learning with Mobilenet. *Journal of Adaptive Learning Technologies*, 1(11), 12-22.
- [5] Betz, F., Oprică, S., Peltonen, T. A., & Sarlin, P. (2014). Predicting distress in European banks. *Journal of Banking & Finance*, 45, 225–241.
- [6] Rahman, M. S., & Kim, E. Y. (2024). Embedding Adaptive Machine-Learning Decision Support in Suspicious Activity aReport Investigation Workflows: An Enterprise Information-Systems Framework. *International Journal of Acta Informatica*, 3(1), 245-258.
- [7] 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.
- [8] Kabir, A. A., Mahmud, F. U., Rahman, M. S., Rashid, S. U., Siddiqui, M. I. H., & Shammah, R. S. (2024). Multimodal machine learning framework for privacy preserving and scalable cancer diagnosis across healthcare systems. *Journal of Adaptive Learning Technologies*, 1(6).

- [9] Chaikin, D. (2009). How effective are suspicious transaction reporting systems? *Journal of Money Laundering Control*, 12(3), 238–253. <https://doi.org/10.1108/13685200910973628>
- [10] Cohn, D. A., Ghahramani, Z., & Jordan, M. I. (1996). Active learning with statistical models. *Journal of Artificial Intelligence Research*, 4, 129–145. <https://doi.org/10.1613/jair.295>
- [11] 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.
- [12] Rashid, S. U., & Cheng, X. (2024). Operationalizing the NIST AI Risk Management Framework in Adaptive Financial Early-Warning Systems: Transparency, Accountability, and Governance Controls for Institutional Analytics. *The Metascience*, 2(4), 142-154.
- [13] 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. <https://doi.org/10.1016/j.ejor.2021.06.023>
- [14] 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. <https://doi.org/10.1016/j.dss.2010.08.006>
- [15] Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R., & Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision Support Systems*, 33(2), 111–126. [https://doi.org/10.1016/S0167-9236\(01\)00139-7](https://doi.org/10.1016/S0167-9236(01)00139-7)
- [16] Takáts, E. (2011). A theory of “crying wolf”: The economics of money laundering enforcement. *Journal of Law, Economics, and Organization*, 27(1), 32–78.
- [17] Rahman, M. S., Islam, C. S., & Tirupathi, D. (2024). Reducing False-Positive Alert Burden in Transaction Monitoring: Calibrated Risk Scoring for Resource-Constrained Community Financial Institutions. *International Journal of Modern Computing*, 7(1), 266-282.
- [18] West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security*, 57, 47–66. <https://doi.org/10.1016/j.cose.2015.09.005>
- [19] Karim, N. B., Vrettos, K., & Seitz, J. (2024a). Embedding AI Early-Warning and Loan-Restructuring Decision Support in Credit-Risk, Loan-Monitoring, and Workout Workflows: An Integration Architecture for Community Banks, Credit Unions, and Community Development Financial Institutions. *The Computertech*, 25-40.

- [20] Arnott, D., & Pervan, G. (2005). A critical analysis of decision support systems research. *Journal of Information Technology*, 20(2), 67–87. <https://doi.org/10.1057/palgrave.jit.2000035>
- [21] Karim, N. B., Pentyala, R. K., & Simko, R. (2024b). 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. *International Journal of Acta Informatica*, 3(1), 259-278.
- [22] Ancker, J. S., Edwards, A., Nosal, S., Hauser, D., Mauer, E., & Kaushal, R. (2017). Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC Medical Informatics and Decision Making*, 17, Article 36.
- [23] Rashid, S. U., Junck, L., & Albous, M. R. (2024a). Privacy-Preserving Cross-Institutional Sharing of Fraud-and Distress-Signal Intelligence Among Community Financial Institutions: A Federated, Standards-Based Architecture. *International Journal of Modern Computing*, 7(1), 283-297.
- [24] 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.
- [25] Rashid, S. U., Siddiqui, M. I. H., Mahmud, F. U., Rahman, M. S., Kabir, A. A., & Shammah, R. S. (2024b). Machine learning based clinical decision support for heart disease prediction using structured patient data. *Journal of Computer Science and Technology Studies*, 6(1), 340-350. (Rashid et al., 2024)
- [26] Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. *AI Magazine*, 35(4), 105–120.
- [27] 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.