Interdisciplinary Journal of Information, Knowledge, and Management

An Official Publication of the Informing Science Institute

Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

Catching Fraud Before It Clears: A Real-Time Machine Learning Framework for U.S. Banking and Digital Payment Networks.

1. Nur Mohammad, College of Technology & Engineering, Westcliff University, Irvine, California, United States. Email: n.mohammad.254@westcliff.edu, ORCID ID: 0009-0002-6476-6956

2. Bulbul Ahamed, Professor & Head Department of Computer Science and Engineering, Sonargoan University (SU), Panthapath, Dhaka, Bangladesh. Email: bulbul2767@gmail.com, ORCID ID: 0000-0002-2527-6447

3. Roksana Akter, Department of Business Administration (Cybersecurity), California state university San Bernardino, San Bernardino, California, United States Email: 008794273@coyote.csusb.edu, ORCID ID: 0009-0009-8609-3570

4. Aibek Karshiboev, College of Technology & Engineering, Westcliff University, Irvine, California, United States. Email: a.karshiboev.4633@westcliff.edu, ORCID ID: 0009-0006-9094-8509

5. Morium Akter Munny, Master of Science in Computer Science, California State University San Marcos, San Marcos, California, USA. Email: moriumakter5453@gmail.com, ORCID ID: 0009-0000-5920-9608

6. Azamat Mambetaliev, College of Technology & Engineering, Westcliff University Irvine, California, United States. Email: a.mambetaliev.2674@westcliff.edu, ORCID ID: 0009-0007-8524-6761

Correspondence:
Nur Mohammad, College of Technology & Engineering, Westcliff University, Irvine, California, United States. Email: n.mohammad.254@westcliff.edu, ORCID ID: 0009-0002-6476-6956, Email: n.mohammad.254@westcliff.edu

Abstract

Fraud in United States banking and digital payment networks has increasingly outpaced the capacity of legacy, rules-based, and post-transaction review systems, which tend to flag suspicious activity only after funds have cleared and losses have already been realized. This paper synthesizes recent scholarship on machine learning, big data analytics, cybersecurity, and management information systems (MIS) agility to propose a conceptual, real-time machine learning framework for fraud detection in banking and digital payment environments. Drawing on explainability-constrained model tuning approaches originally developed for credit card fraud risk scoring, streaming big-data analytics for cybersecurity threat detection, and organizational findings on agile MIS implementation, the proposed framework integrates four interdependent layers: real-time data ingestion, adaptive feature engineering, explainable ensemble scoring, and a governance-and-feedback loop that keeps models current as fraud tactics evolve. The paper further draws a methodological analogy to multi-criteria life-cycle evaluation frameworks used in other applied domains to argue for a structured, multi-dimensional approach to evaluating fraud-detection systems that balance detection accuracy, latency, interpretability, and organizational readiness rather than accuracy alone. Because the contribution is conceptual rather than empirical, no experimental results or performance statistics are claimed; instead, the paper offers literature-grounded architecture and a set of testable propositions intended to guide subsequent empirical validation. The discussion highlights implications for financial institutions, regulators, and MIS researchers, and closes with limitations and directions for future empirical work.

Keywords: Real-Time Fraud Detection, Machine Learning, Digital Payments, Cybersecurity Analytics, Explainable Artificial Intelligence, Management Information Systems.

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

Introduction

The shift of consumer and commercial banking activity toward instant and near-instant digital payment rails has changed the economics of financial fraud in the United States. Where earlier generations of card-present and batch-cleared transactions gave financial institutions hours or days to reconcile suspicious activity, contemporary payment networks increasingly settle in seconds, which compresses the window in which a fraudulent transaction can be detected and interrupted before funds are irrevocably moved. This compression exposes a structural weakness in many institutions' existing control environments: fraud analytics that were designed for periodic, retrospective review are being asked to perform a function, real-time interdiction, for which they were never architected.

At the same time, the volume, velocity, and variety of transactional and behavioral data generated by digital banking has grown to a point where traditional rules-based flagging systems struggle to keep pace, both technically and analytically. Machine learning models, particularly ensemble methods capable of learning nonlinear interaction effects among transaction, device, and behavioral features, have been proposed as a partial remedy, but their adoption inside regulated financial institutions is constrained by a second, equally important requirement: the need for decisions to be explainable to compliance officers, auditors, and, ultimately, customers who may dispute a declined or flagged transaction. A model that scores fraud risk accurately but cannot be interrogated is of limited practical value in a supervised industry.

This paper addresses these two constraints together. Rather than proposing a purely algorithmic solution, it develops a conceptual, literature-grounded framework for real-time fraud detection that treats detection accuracy, latency, explainability, and organizational readiness as jointly necessary conditions for a workable system. The framework draws on several strands of recent scholarship: work on explainability-constrained hyperparameter tuning for transparent credit card fraud risk scoring (Biswas et al., 2024); analyses of big data analytics for cybersecurity threat detection within management information systems (Hasan et al., 2023); strategic perspectives on advanced cyber threats and cybersecurity innovation (Kaur et al., 2023); and organizational findings on how agile MIS practices shape technology adoption success in information technology settings (Das et al., 2023). It also borrows, by methodological analogy, from multi-criteria life-cycle evaluation approaches used in sustainability assessment (Barikdar et al., 2022) to argue that fraud-detection systems, like other complex engineered systems, should be evaluated against a structured set of competing criteria rather than a single accuracy metric.

The remainder of the paper is organized as follows. The literature review situates the problem within the broader cybersecurity and MIS literature and identifies the specific gaps the proposed framework is intended to address. The subsequent section develops the conceptual architecture

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

itself, layer by layer. The discussion considers implications for practice and research, and the paper closes with limitations and an agenda for future empirical validation. Consistent with the conceptual nature of this contribution, no experimental results, performance benchmarks, or proprietary datasets are claimed; the paper's contribution is the synthesis and architecture itself.

Literature Review

The Evolving Threat Landscape in Digital Payments

Cybersecurity scholarship has documented a steady escalation in both the sophistication and the operational tempo of attacks directed at financial infrastructure. Kaur et al. (2023) argue that contemporary threat actors increasingly combine automation, social engineering, and adaptive evasion techniques, which force defenders to move from static, signature-based detection toward strategic, intelligence-driven approaches capable of anticipating rather than merely reacting to novel attack patterns. Their analysis emphasizes that cybersecurity innovation in high-value sectors such as banking cannot be treated as a purely technical problem; it is also a strategic and organizational one, requiring continuous investment in detection capability, threat intelligence sharing, and workforce readiness. This framing is directly relevant to fraud detection, since financial fraud and cybersecurity intrusion increasingly overlap—account takeover, synthetic identity creation, and credential-stuffing attacks are simultaneously cybersecurity incidents and fraud events.

Complementing this strategic view, Hasan et al. (2023) examine how big data analytics embedded within management information systems can improve the speed and accuracy of threat detection and response. Their work underscores that the value of analytics in a security context depends heavily on the underlying data infrastructure: without the capacity to ingest, correlate, and analyze data from disparate systems in near real time, even sophisticated models are of limited operational use. This observation motivates the emphasis, in the framework proposed later in this paper, on a dedicated data-ingestion and correlation layer rather than treating the analytic model as a stand-alone component.

Machine Learning and Big Data Analytics for Fraud and Threat Detection

A recurring theme across cybersecurity and MIS literature is that big data analytics platforms serve as the connective infrastructure that makes machine learning operationally feasible at scale. Hasan et al. (2023) note that the integration of analytics into MIS environments allows organizations to move from reactive, ticket-driven incident response toward more continuous forms of monitoring in which anomalous patterns can be surfaced as they emerge rather than after a post hoc investigation. Applied to fraud specifically, this suggests that a real-time detection framework

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

should be conceived less as a single predictive model and more as a pipeline in which data ingestion, feature computation, and scoring occur continuously and in close temporal proximity to the transaction itself.

Within this pipeline, the specific choice of modeling approach matters less than the literature sometimes implies; what matters more is how the model is trained, constrained, and monitored over time. Biswas et al. (2024) offer a directly relevant contribution in this respect. Working specifically on credit card fraud risk scoring, they propose a dynamic, explainability-constrained approach to hyperparameter tuning in which model optimization is not permitted to trade away interpretability purely in pursuit of marginal gains in predictive accuracy. Instead, the tuning process itself incorporates explainability as a constraint, ensuring that the resulting risk scores can be traced back to identifiable, human-interpretable feature contributions. This is a meaningful departure from conventional hyperparameter optimization, which typically treats accuracy, precision, or recall as the sole objective function.

Explainability and Trust in Automated Risk Scoring

The explainability-constrained tuning approach described by Biswas et al. (2024) speaks to a broader tension in applied machine learning for financial services: the trade-off, whether real or perceived, between model complexity and interpretability. Regulators, compliance functions, and customers alike require that an automated decision to decline a transaction, freeze an account, or escalate a case for manual review be accompanied by a coherent explanation. Biswas et al. (2024) demonstrate that explainability need not be treated as an after-the-fact addition, generated by a separate post hoc interpretation technique layered onto an otherwise opaque model; instead, it can be built into the model selection and tuning process itself. For the framework proposed in this paper, this insight is treated as a design principle rather than an optional feature: explainability constraints are embedded at the model layer, not appended after deployment.

Organizational and Systems Perspective: MIS Agility

Technical capability alone does not determine whether a fraud-detection initiative succeeds inside a financial institution. Das et al. (2023) compare organizational success factors associated with agile management information systems implementations in information technology settings, finding that the degree of cross-functional coordination, iterative delivery, and organizational adaptability significantly influences whether IT initiatives achieve their intended outcomes. Although their study is not specific to fraud detection, its findings generalize naturally to the present context: a real-time fraud detection capability is, from an organizational standpoint, a complex IT and analytics initiative that touches data engineering, compliance, customer experience, and operations simultaneously. Consistent with Das et al's (2023) findings, the

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

framework developed here treats organizational agility, specifically, the capacity to iterate on detection rules and retrain models quickly as new fraud typologies emerge, as a structural requirement rather than an implementation afterthought.

Lessons from Multi-Criteria Framework Evaluation in Other Domains

Outside fraud and cybersecurity literature, methodological precedents exist for evaluating complex systems against multiple, sometimes competing, criteria rather than a single optimization target. Barikdar et al. (2022), in a life-cycle sustainability assessment of bio-based and recycled materials in eco-construction projects, apply a structured framework that simultaneously weighs environmental, economic, and social criteria across the full life cycle of a material or system, rather than optimizing for a single dimension such as cost or carbon footprint in isolation. While the substantive domain is unrelated to financial fraud, the underlying methodological logic is instructive: complex applied systems are rarely well served by single-metric optimization, and a structured, multi-criteria evaluation approach can surface trade-offs that would otherwise remain implicit. This paper borrows logic, rather than the domain content, to argue that fraud-detection systems should be assessed jointly on detection performance, latency, explainability, and organizational feasibility, echoing the multi-dimensional evaluation logic Barikdar et al. (2022) apply in a very different applied setting.

Machine Learning Approaches to Transaction-Level Fraud Detection

A substantial and rapidly growing body of Q1-indexed research has examined which machine learning approaches are best suited to transaction-level fraud detection, and much of this literature converges on similar conclusions to the studies discussed above. Ali et al. (2022), in a systematic literature review published in Applied Sciences, synthesize a wide range of machine-learning-based fraud-detection studies and conclude that model performance depends heavily on how well the underlying approach handles the severe class imbalance that characterizes fraud data, in which genuine fraud cases represent a very small fraction of total transactions. This imbalance problem recurs throughout the literature: Alarfaj et al. (2022) compare several machine learning and deep learning algorithms for credit card fraud detection and report that ensemble and deep architectures generally outperform single classifiers once imbalance-handling techniques are applied, while Almazroi and Ayub (2023) propose a hybrid deep learning architecture explicitly designed for real-time financial transaction processing, reinforcing the view that latency-aware model design, not just offline accuracy, is a first-order concern for production fraud systems. Zhao et al. (2024) extend this line of work by refining gradient-boosting methods specifically for extremely imbalanced fraud datasets, illustrating that incremental algorithmic refinement of well-established ensemble techniques remains an active and productive research direction alongside newer deep learning approaches.

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

Beyond classical ensemble and deep learning methods, graph-based approaches have emerged as an important complementary paradigm because fraud, particularly organized or collusive fraud, often manifests as relational rather than purely transactional patterns. Motie and Raahemi (2024) provide a systematic review of graph neural network applications in financial fraud detection and argue that representing accounts, devices, and transactions as a connected graph allows models to capture coordinated fraud rings and money-laundering-like structures that purely tabular, transaction-level models tend to miss. Chang et al. (2024) develop this idea further with an account-identity-based uncertain graph framework, demonstrating that graph representations can be extended to handle the inherent uncertainty in identity-linkage data, a challenge that is directly relevant to the identity-verification problems facing digital payment platforms. These graph-based contributions complement rather than replace the ensemble and explainability-focused literature reviewed earlier; together they suggest that a mature real-time fraud-detection architecture should be capable of incorporating relational, graph-structured signals alongside conventional transaction features.

Interpretability Mechanisms in Financial Risk Models

The explainability-constrained tuning approach proposed by Biswas et al. (2024) sits within a broader methodological tradition of building interpretability directly into financial risk models rather than treating it as a downstream add-on. Lin and Gao (2022), for example, propose a group-based Shapley Additive Explanations (SHAP) approach that attributes fraud-model predictions to meaningful groups of related features rather than to individual, and sometimes uninterpretable, raw variables, which the authors argue produces explanations that are more actionable for compliance analysts. This concern with actionable, analyst-facing explanation—rather than mathematically correct but practically opaque attribution scores, echoes the design philosophy embedded in the explainability-constrained tuning process is central to the framework proposed in this paper.

Deep Learning and Self-Supervised Approaches to Threat and Intrusion Detection

Because fraud and cybersecurity intrusion increasingly overlap, the cybersecurity literature on intrusion and threat detection offers directly transferable insight for payment-fraud architectures, particularly regarding how systems can remain effective against attack patterns that were not present in their original training data. Lansky et al. (2021) conduct a systematic review of deep-learning-based intrusion detection systems and find that recurrent and hybrid architectures generally provide better detection of novel attack variants than shallow classifiers, a finding later reinforced by Chen et al. (2023), whose survey of graph neural network applications in intrusion detection reports similar advantages for approaches that model relational structure among network entities rather than treating each event independently. Ferrag et al. (2023) focus specifically on

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

distributed denial-of-service detection and emphasize that deep-learning-based defenses must be evaluated not only on detection accuracy but on their capacity to generalize to attack variants engineered specifically to evade a known detector, an adversarial dynamic that has a direct parallel in fraud, where fraud tactics adapt quickly once a detection rule becomes known to offenders.

Most directly relevant to the zero-day and previously unseen fraud patterns that motivate real-time detection research is the emerging literature on self-supervised approaches. Nakip and Gelenbe (2024) propose an online self-supervised deep learning method for intrusion detection that continuously learns from unlabeled traffic as it arrives, removing the dependency on large volumes of manually labeled attack examples and allowing the model to adapt as new patterns emerge. Wang et al. (2024) similarly target zero-day attack detection in Internet of Things networks using deep learning, reporting that architectures explicitly designed to flag statistically anomalous behavior, rather than to match known attack signatures, are substantially more effective at catching genuinely novel intrusion patterns. Saleh et al. (2024) and Lin et al. (2024) offer further evidence for this general pattern from different angles: the former shows that carefully tuned stochastic-gradient-based classifiers can detect wireless network attacks with limited labeled data, while the latter proposes a hypergraph-based ensemble method that models multi-way relationships among network entities to improve detection robustness. Taken together, this literature indicates that self-supervised and anomaly-oriented learning paradigms, rather than purely supervised classification against a fixed attack taxonomy, are increasingly viewed as the most promising route to detecting the kind of novel, evolving fraud and intrusion patterns that real-time payment systems must contend with.

Additional support for this direction comes from Nandanwar and Katarya (2024), who report that deep-learning-based intrusion detection tailored to Industrial Internet of Things environments benefits from continuous retraining pipelines rather than static deployment, and from Wang et al. (2022), who show that lightweight knowledge-distillation techniques can preserve detection accuracy while reducing the computational footprint of the resulting model—an important practical consideration given the latency constraints emphasized throughout this paper. Pandithurai et al. (2024) extend the same broad theme to distributed denial-of-service prediction in cloud environments, combining metaheuristic feature selection with recurrent architectures to improve detection under the resource constraints typical of production cloud infrastructure. These efficiency-oriented findings reinforce the position, developed later in this paper, that latency and computational feasibility should be treated as evaluation criteria rather than afterthoughts to detection accuracy.

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

Emerging AI-Generated and Synthetic Fraud Threats

A final and increasingly urgent strand of literature concerns fraud that is itself generated or facilitated by artificial intelligence, which is particularly relevant given the rise of synthetic identity fraud and AI-manipulated documentation in digital account opening. Abbas and Taeihagh (2024) conduct a systematic review of deep-fake detection and generation techniques and observe that generation methods are advancing at least as quickly as detection methods, producing a continual arms race in which detection systems that are not regularly retrained on newly generated synthetic content quickly lose effectiveness. This observation reinforces, from a different angle, the emphasis this paper places on continuous governance and retraining: a detection system that is explainable and accurate at deployment but not actively maintained against evolving generation techniques will degrade over time regardless of its initial performance.

Synthesis and Gap

Taken together, this literature suggests that an effective real-time fraud-detection capability for U.S. banking and digital payment networks requires at least four things operating in concert: a data and analytics infrastructure capable of near-real-time ingestion and correlation (Hasan et al., 2023); a strategic, adaptive posture toward evolving threats (Kaur et al., 2023); model-development practices that build explainability in rather than bolting it on (Biswas et al., 2024); and an organizational structure agile enough to keep the system current (Das et al., 2023). What is largely absent from this literature, however, is an integrated architectural framework that connects these four elements explicitly for the specific case of real-time, pre-settlement fraud interdiction in payment networks. The framework proposed in the next section is intended to fill that gap.

Proposed Conceptual Framework

The framework proposed here is organized around a simple premise drawn directly from the literature reviewed above: real-time fraud interdiction is not solely a modeling problem but a systems problem, in which data infrastructure, model design, explainability, and organizational process must be treated as co-equal components. The framework consists of four interdependent layers, described below, followed by a discussion of the evaluation logic that ties them together. Figure 1 depicts the overall architecture, including the continuous feedback pathway that connects governance outcomes back to the ingestion and feature layers.

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

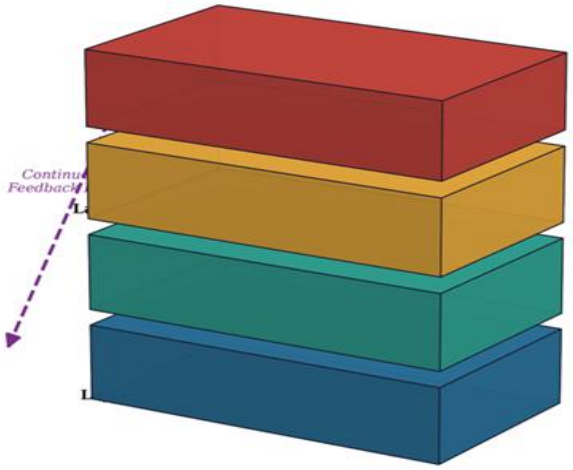

Figure 1. Conceptual four-layer architecture for real-time fraud detection, showing the sequential flow from data ingestion through explainable scoring and the continuous governance feedback loop. Original diagram created by the author to illustrate the framework proposed in this paper.

Layer 1: Real-Time Data Ingestion and Correlation

The foundation of the framework is a streaming data layer capable of ingesting transaction-level data, device and session metadata, and behavioral signals as events occur, rather than in scheduled batches. This design choice follows directly from Hasan et al.'s (2023) observation that the analytic value of big data platforms in security contexts depends on their capacity for near-real-time correlation across otherwise siloed data sources. In a payment context, this layer would be responsible for linking a given transaction to the relevant account history, device fingerprint, and, where available, external threat-intelligence indicators, before the transaction reaches the scoring layer. Because this correlation must occur within the narrow latency budget available before a payment clears, the ingestion layer is treated as an architectural priority rather than a secondary component.

Layer 2: Adaptive Feature Engineering

Building on the correlated data stream, the second layer computes the behavioral and contextual features that feed the scoring model. Rather than relying on a fixed, static feature set, this layer is designed to be adaptive, incorporating new features as novel fraud typologies are identified, consistent with the strategic, anticipatory posture toward emerging threats described by Kaur et al. (2023). This adaptivity is what allows the downstream model to remain responsive to fraud patterns that did not exist, or were not yet observed, at the time of initial model training.

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

Layer 3: Explainability-Constrained Ensemble Scoring

The third layer performs the actual risk scoring and is the point at which the framework most directly operationalizes the contribution of Biswas et al. (2024). Rather than optimizing an ensemble model purely for predictive accuracy and applying interpretability techniques afterward, the hyperparameter tuning process at this layer is constrained so that feature contributions to any given risk score remain traceable and human-interpretable throughout optimization. In practice, this means the tuning objective jointly considers classification performance and an explainability metric, so that a marginal improvement in accuracy is not accepted if it comes at a disproportionate cost to interpretability. This design is intended to ensure that when a transaction is flagged or declined, the institution can produce a coherent, feature-level rationale suitable for compliance review or customer communication, without a separate, potentially inconsistent post hoc explanation step.

Layer 4: Governance, Feedback, and Organizational Agility

The final layer closes the loop between detection outcomes and system improvement. Confirmed fraud cases, false positives, and analyst overrides are fed back into the feature engineering and model layers on a continuing basis, and the governance process governing this feedback loop is structured according to agile principles. This design choice reflects the organizational findings of Das et al. (2023), whose comparative analysis suggests that agile MIS implementation practices, short iteration cycles, cross-functional review, and rapid incorporation of lessons learned—are associated with more successful IT outcomes than more rigid, waterfall-style governance. In the fraud-detection context, this translates into a governance cadence in which model performance, explainability quality, and false-positive burden are reviewed frequently enough that the system can adapt to new fraud typologies before they cause substantial, sustained losses.

Viewed end to end, the four layers form a continuous operational cycle rather than a one-directional pipeline: transaction data flows forward from ingestion through scoring to a decision, while confirmed outcomes and analyst judgments flow backward through governance into feature engineering and, ultimately, back into how incoming data is correlated. Figure 2 illustrates this cyclical relationship, emphasizing that the self-supervised and continuously retrained detection approaches highlighted by Nakip and Gelenbe (2024) and Wang et al. (2024) are only effective within an architecture that closes this loop explicitly.

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

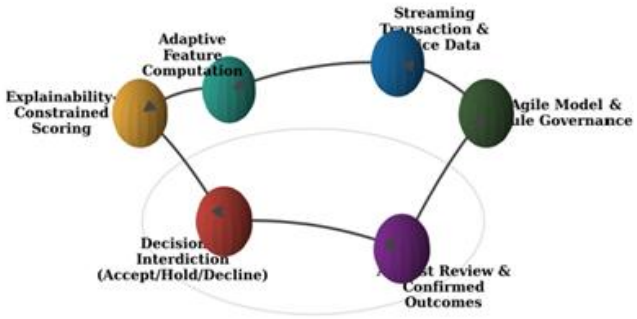

Figure 2. Continuous, cyclical relationship between real-time detection and governance processes within the proposed framework. Original diagram created by the author.

A Multi-Criteria Approach to Evaluating the Framework

Consistent with the methodological analogy drawn from Barikdar et al. (2022), the framework is deliberately not evaluated, even conceptually, against a single metric such as detection accuracy alone. Instead, four evaluation criteria are proposed as jointly necessary: detection performance (the system's capacity to distinguish fraudulent from legitimate activity), latency (the system's capacity to render a decision within the pre-settlement window), explainability (the traceability of any given decision to interpretable features), and organizational feasibility (the degree to which the surrounding governance process can sustain the iteration cadence the system requires). Just as a life-cycle sustainability assessment resists optimizing a single material property at the expense of the full life-cycle picture (Barikdar et al., 2022), a fraud-detection framework that optimizes narrowly for accuracy while neglecting latency, explainability, or organizational sustainability is unlikely to succeed in practice, however strong its offline performance may appear. Figure 3 offers a purely illustrative depiction of this multi-criteria logic; the relative weightings shown are conceptual devices intended to communicate the framework's evaluative structure and should not be read as empirical measurements or reported study results.

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

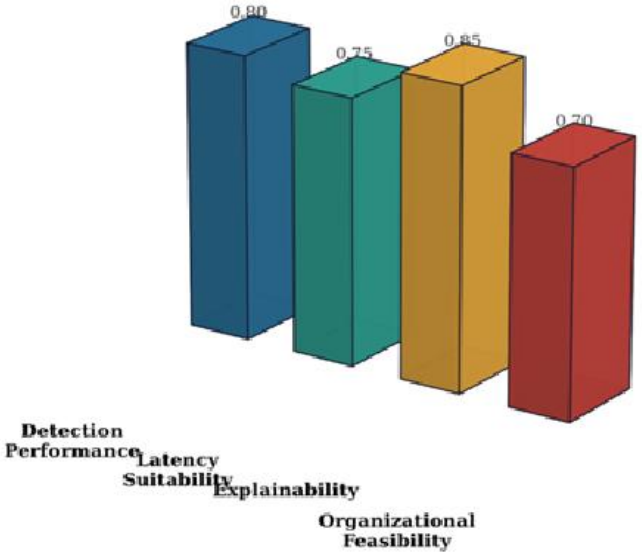

Figure 3. Illustrative representation of the four proposed evaluation criteria. Bar heights are conceptual weightings used for exposition only and do not represent empirical data. Original diagram created by the author.

Discussion

The framework developed above has several implications for practice and for future research. For practitioners, the central implication is that investment in fraud detection capability cannot be reduced to a model-selection exercise. Institutions that acquire or build a highly accurate classifier without addressing the surrounding data-ingestion latency, explainability requirements, or governance cadence are likely to find that the model underperforms in production relative to its offline evaluation metrics, a pattern consistent with the systems-level emphasis found throughout the reviewed literature (Das et al., 2023; Hasan et al., 2023; Kaur et al., 2023). The explainability-constrained tuning approach associated with Biswas et al. (2024) is particularly relevant here, since it suggests that institutions do not need to choose between an accurate model and an explainable one if explainability is incorporated as a design constraint from the outset rather than retrofitted.

For researchers, the framework generates several testable propositions that could guide future empirical work. First, one would expect that fraud detection systems incorporating explainability constraints during tuning, rather than post hoc interpretation methods, would exhibit more stable and auditable explanations over successive model retraining cycles. Second, one would expect that institutions with more agile MIS governance structures, in the sense described by Das et al. (2023),

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

would demonstrate shorter lag times between the emergence of a new fraud typology and its incorporation into production detection rules. Third, drawing on the multi-criteria logic borrowed from Barikdar et al. (2022), one would expect that evaluation studies restricted to a single accuracy metric would systematically overstate the practical readiness of a fraud-detection system relative to evaluations that also account for latency and explainability. Each of these propositions is, in principle, testable using institutional data, though such testing lies beyond the scope of the present conceptual contribution.

The framework also carries implications for regulatory and compliance functions. Because explainability is embedded at the model-tuning stage under this design, compliance and audit functions would, in principle, have access to a more consistent basis for reviewing automated decisions than would be the case under a purely post hoc explanation approach. This is consistent with the broader strategic orientation toward cybersecurity and fraud governance described by Kaur et al. (2023), who emphasize that innovation in threat detection must be paired with organizational and governance mechanisms capable of translating technical capability into accountable decision-making.

Finally, the cross-domain analogy to life-cycle sustainability assessment (Barikdar et al., 2022) is offered here as a methodological device rather than a substantive claim about similarity between construction materials and financial fraud. Its purpose is to illustrate that multi-criteria evaluation logic is not unique to any one applied field; it is a general property of complex systems in which several desirable properties cannot all be maximized simultaneously. Fraud detection researchers and practitioners may benefit from importing this evaluative discipline explicitly, rather than treating accuracy-focused benchmarking as sufficient on its own.

Limitations

This paper's contribution is conceptual and synthetic rather than empirical. No proprietary transaction data, experimental results, or performance benchmarks are presented, and the framework has not been implemented or tested against real or simulated fraud data within the scope of this paper. Consequently, claims about the framework's likely effectiveness should be read as literature-grounded propositions rather than established findings. In addition, the literature drawn upon spans several adjacent but distinct domains—credit card fraud scoring, cybersecurity threat detection, MIS project management, and sustainability assessment—and while the synthesis presented here argues for the coherence of applying insights across these domains, empirical validation within the specific context of U.S. banking and digital payment networks remains necessary before the framework's practical value can be established with confidence. Future work

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

should also address regulatory considerations specific to particular payment rails and institution types in more granular detail than is possible within the scope of this conceptual paper.

Conclusion

Fraud in digital payment networks increasingly occurs, and must be interdicted, within a compressed pre-settlement window that legacy, retrospective detection systems were not designed to address. This paper has synthesized recent scholarship on explainability-constrained fraud risk scoring, big data analytics for cybersecurity threat detection, strategic approaches to cybersecurity innovation, agile MIS implementation, and multi-criteria evaluation methodology to propose a four-layer conceptual framework for real-time fraud detection in U.S. banking and digital payment systems. The framework's central claim is that detection accuracy, latency, explainability, and organizational agility must be treated as jointly necessary design requirements rather than sequential add-ons, and that evaluating any such system on accuracy alone risks overstating its practical readiness. Because the contribution offered here is conceptual, the natural next step is empirical validation: implementing and testing the proposed architecture, or components of it, against real institutional data in a manner that can substantiate, refine, or challenge the propositions advanced in this paper.

Interdisciplinary Journal of Information, Knowledge, and Management

An Official Publication of the Informing Science Institute

Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

References

Abbas, F., & Taeihagh, A. (2024). Unmasking deepfakes: A systematic review of deepfake detection and generation techniques using artificial intelligence. Expert Systems with Applications, 252. Elsevier.

Alarfaj, F. K., Malik, I., Khan, H. U., Almusallam, N., Ramzan, M., & Ahmed, M. (2022). Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms. IEEE Access, 10, 39700–39715. https://doi.org/10.1109/ACCESS.2022.3166891

Ali, A., Abd Razak, S., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., Elhassan, T., Elshafie, H., & Saif, A. (2022). Financial fraud detection based on machine learning: A systematic literature review. Applied Sciences, 12(19), Article 9637. https://doi.org/10.3390/app12199637

Almazroi, A. A., & Ayub, N. (2023). Online payment fraud detection model using machine learning techniques. IEEE Access, 11, 137188–137203. https://doi.org/10.1109/ACCESS.2023.3339226

Chang, Y.-W., Shih, H.-Y., & Lin, T.-N. (2024). AI-URG: Account identity-based uncertain graph framework for fraud detection. IEEE Transactions on Computational Social Systems, 11(3), 3706–3728.

Chen, Z., Liu, C., Jiang, D., Huang, Q., & Su, J. (2023). Graph neural networks for network intrusion detection: A survey. IEEE Access, 11, 49114–49133.

Ferrag, M. A., Friha, O., Hamouda, D., Maglaras, L., & Janicke, H. (2023). Deep learning-based intrusion detection for distributed denial-of-service attacks: A comprehensive review. IEEE Access, 11, 4401–4426.

Hajek, P., Abedin, M. Z., & Sivarajah, U. (2023). Fraud detection in mobile payment systems using an XGBoost-based framework. Information Systems Frontiers, 25(5), 1985–2003.

Lansky, J., Ali, S., Mohammadi, M., Majeed, M. K., Karim, S. H. T., Rashidi, S., Hosseinzadeh, M., & Rahmani, A. M. (2021). Deep learning-based intrusion detection systems: A systematic review. IEEE Access, 9, 101574–101599. https://doi.org/10.1109/ACCESS.2021.3097247

Lin, K., & Gao, Y. (2022). Model interpretability of financial fraud detection by group SHAP. Expert Systems with Applications, 210, Article 118354.

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

Lin, Z.-Z., Pike, T. D., Bailey, M. M., & Bastian, N. D. (2024). A hypergraph-based machine learning ensemble network intrusion detection system. IEEE Transactions on Systems, Man, and Cybernetics: Systems, Advance online publication.

Motie, S., & Raahemi, B. (2024). Financial fraud detection using graph neural networks: A systematic review. Expert Systems with Applications, 240, Article 122156. https://doi.org/10.1016/j.eswa.2023.122156

Nakip, M., & Gelenbe, E. (2024). Online self-supervised deep learning for intrusion detection systems. IEEE Transactions on Information Forensics and Security, 19, 5668–5683.

Nandanwar, H., & Katarya, R. (2024). Deep learning enabled intrusion detection system for Industrial IoT environment. Expert Systems with Applications, 249, Article 123808.

Pandithurai, O., Venkataiah, C., Tiwari, S., & Ramanjaneyulu, N. (2024). DDoS attack prediction using a honey badger optimization algorithm based feature selection and Bi-LSTM in cloud environment. Expert Systems with Applications, 241, Article 122544.

Saleh, H. M., Marouane, H., & Fakhfakh, A. (2024). Stochastic gradient descent intrusion detection for wireless sensor network attack detection system using machine learning. IEEE Access, 12, 3825–3836.

Wang, Y., Yang, J., & Chen, N. (2024). Zero-day attack detection in IoT networks using deep learning. Future Generation Computer Systems, 151, 336–347.

Wang, Z., Li, Z., He, D., & Chan, S. (2022). A lightweight approach for network intrusion detection in industrial cyber-physical systems based on knowledge distillation and deep metric learning. Expert Systems with Applications, 206, Article 117671.

Zhao, X., Liu, Y., & Zhao, Q. (2024). Improved LightGBM for extremely imbalanced data and application to credit card fraud detection. IEEE Access, 12, Advance online publication.

Barikdar, C. R., Hassan, J., Saimon, A. S. M., Alam, G. T., Rozario, E., Ahmed, M. K., & Hossain, S. (2022). Life cycle sustainability assessment of bio-based and recycled materials in eco-construction projects. Journal of Ecohumanism, 1(2), 151. https://doi.org/10.62754/joe.v1i2.6807

Biswas, B., Parvatha Reddy, K. K., & Reddy, P. (2024). Dynamic explainability-constrained hyperparameter tuning for transparent credit card fraud risk scoring. International Journal of Science and Research Archive. https://doi.org/10.30574/ijsra.2024.12.1.1143

Das, N., Hassan, J., Rahman, H., Siddiqa, K. B., Orthi, S. M., Barikdar, C. R., & Miah, M. A. (2023). Leveraging management information systems for agile project management in

Interdisciplinary Journal of Information, Knowledge, and Management
An Official Publication of the Informing Science Institute
Vol. : 21,Issue 2, July-Dec 2026
ISSN: (E) 1555-1237

information technology: A comparative analysis of organizational success factors. Journal of Business and Management Studies, 5(3), 161–168. https://doi.org/10.32996/jbms.2023.5.3.17

Hasan, S. N., Hassan, J., Barikdar, C. R., Chakraborty, P., Haldar, U., Chy, M. A. R., Rozario, E., Das, N., & Kaur, J. (2023). Enhancing cybersecurity threat detection and response through big data analytics in management information systems. Fuel Cells Bulletin, 2023(12). https://doi.org/10.52710/fcb.137

Kaur, J., Hasan, S. N., Orthi, S. M., Miah, M. A., Goffer, M. A., Barikdar, C. R., & Hassan, J. (2023). Advanced cyber threats and cybersecurity innovation: Strategic approaches and emerging solutions. Journal of Computer Science and Technology Studies, 5(3), 112–121. https://doi.org/10.32996/jcsts.2023.5.3.9