DOI: https://doi.org/10.32782/2524-0072/D2026-86-138

УДК 338.24:004.056.5:519.8

A COST-SENSITIVE ANOMALY DETECTION FRAMEWORK FOR ECONOMIC INFORMATION SECURITY MANAGEMENT

МЕТОДОЛОГІЧНИЙ ПІДХІД ДО ВИЯВЛЕННЯ АНОМАЛІЙ З УРАХУВАННЯМ АСИМЕТРІЇ ВИТРАТ В УПРАВЛІННІ ЕКОНОМІЧНОЮ ТА ІНФОРМАЦІЙНОЮ БЕЗПЕКОЮ ПІДПРИЄМСТВА

Derbentsev Vasyl
PhD in Economics, Professor of the Department of Informatics and Systemology,
Kyiv National Economic University named after Vadym Hetman
ORCID: https://orcid.org/0000-0002-8988-2526

Kroshko Ivan
PhD student of the Department of Computer Science,
Kyiv National Economic University named after Vadym Hetman
ORCID: https://orcid.org/0009-0006-3680-4029

Дербенцев Василь Джоржович, Крошко Іван Андрійович
Київський національний економічний університет імені Вадима Гетьмана

The rapid digitalization of business environments has made information security a critical component of enterprise risk management. This paper develops a formal framework for anomaly detection as an instrument of economic information security management. The business information space is formalized as a tri-domain environment integrating transactional, behavioral, and system-level data. A hybrid detection architecture combines an unsupervised ensemble of Isolation Forest, Autoencoder, and One-Class SVM with a supervised classifier, enabling detection of both novel and known fraud patterns. A cost-sensitive loss function with analytical threshold optimization aligns detection decisions with expected financial loss minimization. A four-tier risk decision matrix translates detection outputs into economically grounded management responses.

Keywords: information and economic security management; anomaly detection; economic risk assessment; multi-source detection architecture; cost-sensitive optimization; hybrid machine learning; digital business.

Цифрова трансформація економічних систем суттєво змінила ландшафт інформаційної безпеки сучасних підприємств. Традиційні підходи, що базуються на жорстко заданих правилах і сигнатурах, виявляються структурно неадекватними: вони здатні виявляти лише заздалегідь каталогізовані загрози і не забезпечують економічної інтерпретації виявлених подій. Метою статті є розробка формального методологічного підходу до виявлення аномалій як інструменту управління економічною та інформаційною безпекою підприємства. Запропонований підхід формалізує інформаційний простір бізнесу як три-доменне середовище, що охоплює транзакційні, поведінкові та системно-рівневі дані, та вводить уніфікований багатоджерельний показник аномальності на основі їх зваженої агрегації. Головною інновацією є гібридна архітектура системи детекції, що інтегрує ансамбль методів машинного навчання без вчителя (Isolation Forest, Autoencoder та One-Class SVM) для виявлення нових загроз із контрольованими моделями класифікації для детекції відомих патернів шахрайства. Вагові коефіцієнти агрегації доменних оцінок оптимізуються відносно економічної функції втрат, а не статистичних метрик точності. Поріг класифікації переосмислюється як економічна змінна рішення, що мінімізує сукупні очікувані фінансові втрати. Функція втрат враховує асиметрію витрат помилкових спрацювань та пропущених загроз, відношення яких у контексті електронної комерції перевищує 7:1. Практична цінність полягає в усуненні трьох структурних обмежень наявних методів: доменної фрагментації аналізу, відсутності економічної інтерпретації результатів детекції та калібрування порогів без урахування асиметрії витрат помилок класифікації. Запропонований підхід забезпечує менеджерам можливість оцінювати ефективність системи безпеки у фінансових показниках замість непрозорих статистичних метрик. Додатково запропонована модель забезпечує адаптивне оновлення параметрів детекції відповідно до динаміки бізнес-процесів та змін цифрового середовища підприємства.

Ключові слова: управління інформаційною та економічною безпекою; виявлення аномалій; оцінка економічних ризиків; мультиджерельна архітектура детекції; оптимізація, чутлива до витрат; гібридне машинне навчання; цифровий бізнес.

Problem statement. The rapid digitalization of economic systems has transformed the structure and risk profile of modern business environments. Enterprises increasingly operate within integrated information ecosystems that combine transactional data, user behavior, and system-level events. While such integration enhances efficiency and scalability, it also increases exposure to complex and evolving security threats, making information security a key component of enterprise risk management with direct financial implications. According to industry data, the average global cost of a data breach has reached record levels [1], while annual losses from e-commerce fraud exceed $48 billion [2].

Traditional security mechanisms based on rule-based and signature-driven detection are increasingly inadequate in this context. Their reliance on predefined threat patterns limits the ability to identify novel or adaptive attacks, which account for a substantial share of security losses in digital commerce and financial systems.

Anomaly detection offers an alternative approach by identifying deviations from normal behavior rather than matching known patterns. However, despite advances in machine learning, its application in business security remains constrained by three limitations. First, existing methods are typically domain-specific and fail to capture cross-domain threat patterns. Second, detection outputs are expressed in statistical terms without direct economic interpretation, limiting their usefulness for managerial decision-making. Third, model calibration is commonly based on statistical metrics that assume symmetric misclassification costs, which contradicts real-world conditions where missed threats are significantly more costly than false alarms.

To address these limitations, this study develops a formal framework for anomaly detection as an instrument of economic information security management. The approach integrates a multi-domain representation of the business information space, a unified anomaly scoring mechanism with economically calibrated weights, and a cost-sensitive optimization layer that aligns detection decisions with expected loss minimization.

In addition, the framework incorporates a hybrid architecture combining unsupervised and supervised methods, enabling the detection of both novel anomalies and known fraud patterns. An economic risk score is introduced to link detection outputs with financial impact, supporting risk-based prioritization of decisions.

The main contribution lies in integrating anomaly detection into a unified economic decision-making framework that connects data representation, model design, and decision rules through the principle of expected loss minimization [3].

Analysis of recent research and publications. In the foundational study by Chandola V. et al. [4], a widely adopted taxonomy was established, distinguishing between point, contextual, and collective anomalies, each associated with distinct detection mechanisms and economic implications in the case of non-detection. Aggarwal C. [5] demonstrated that in high-dimensional and heterogeneous data spaces no single algorithm consistently outperforms others across all anomaly types. This is empirically confirmed in [6], where comparative evaluation of fourteen unsupervised algorithms showed that structurally different families exhibit complementary strengths – the basis for ensemble construction. Authors of [7] identify multi-source data integration and cost-sensitive model calibration as the two most underexplored directions in recent research on anomaly detection.

In the domain of fraud detection, Baesens B. et al. [8] provide the foundational analytical framework across supervised, unsupervised, and network-based approaches. Hilal et al. (2022) show that purely supervised models are inherently unable to identify novel schemes, driving convergence toward hybrid architectures. In the e-commerce context, Rodrigues V. et al. [10], reviewing 64 empirical studies, find that gradient boosting consistently achieves state-of-the-art performance for transaction-level classification, while identifying multi-source signal integration as the most significant unresolved challenge. Authors of [11] demonstrate that rule-based systems achieve only limited effectiveness (52-58% detection rate) for novel fraud, while machine learning ensembles substantially improve detection rates (87-94%). The inclusion of behavioral and profile-based features yields measurable improvements in classification performance, as confirmed by Byrapu Reddy S. et al. [12] and Zeng Q. et al. [13].

In the cybersecurity domain, authors of [14] establish that behavioral baselines in intrusion detection systems are inherently non-stationary and require continuously adapting models. Ahmed M. et al. [15] document that persistently high false positive rates (40-70% in deployed

systems) remain the primary operational barrier to practical adoption. Maci F. et al. [16] demonstrate the effectiveness of unsupervised UEBA for unknown anomalous behavior, with Isolation Forest [17] and One-Class SVM [18] serving as complementary algorithmic foundations.

From an economic perspective, Gordon L. and Loeb M. [3] establish that optimal security investment is governed by expected loss rather than technical performance, implying that detection thresholds should reflect financial consequences of classification errors. IBM Security [1] and Cybersource [2] underscore the growing economic stakes. Höppner S. et al. [19] formalize this at the model level, showing that cost-sensitive calibration yields significant economic gains over accuracy-based approaches and that the optimal threshold derives analytically from relative error costs. The present study extends this principle to the full detection architecture.

Highlighting previously unresolved parts of the overall problem. Despite these advances, the literature remains fragmented along three key dimensions: the separation of data domains, the lack of economic interpretation of anomaly scores, and the absence of unified cost-sensitive optimization frameworks. These limitations motivate the development of an integrated approach that combines multi-source data, hybrid detection models, and economically grounded decision rules within a single analytical framework.

The purpose of the article is to develop an economically grounded anomaly detection framework that integrates multi-source data, hybrid modeling, and cost-sensitive decision-making.

Summary of the main research material. The proposed framework introduces a formal representation of the business information space in which observations are modeled as multi-domain feature vectors integrating transactional, behavioral, and system-level data. This representation reflects the structural complexity of modern digital business environments, where economic activity is generated through the interaction of heterogeneous processes operating at different levels of abstraction. Unlike traditional approaches that treat these data sources independently, the proposed model explicitly accounts for their joint structure, enabling the detection of anomalies that emerge only through cross-domain interactions.

The framework is developed in several stages. First, we formalize the multi-domain information space. Second, we define anomaly detection functions for each domain. Third, we integrate domain-specific scores. Fourth, we establish cost-sensitive decision rules.

Let the set of all observations generated by the information system be:

$$X = \{x_1, x_2, \ldots, x_n\},\ x_i \in \mathbb{R}^m \qquad (1)$$

where n is the total number of observations; m – is the feature space dimensionality.

Each observation may carry a latent binary label $y_i \in \{0,1\}$, where $y_i = 1$ denotes a fraudulent or anomalous event. In practice these labels are frequently incomplete or delayed: the primary motivation for including unsupervised detection mechanisms.

They correspond to a recorded event within the business system. Such events may include financial transactions, user actions, or system-level processes. The high dimensionality of the feature space reflects the richness of available information, but also introduces challenges related to noise, redundancy, and non-stationarity.

To address these challenges, each observation is decomposed into domain-specific components as:

$$X = \{X^{trans}, X^{beh}, X^{\inf o}\}, \qquad (2)$$

where X^{trans} – captures financial transaction parameters (amount, payment method, address correspondence, geographic origin, temporal attributes); X^{beh} – reflects user interaction patterns over time (session velocity, navigation sequences, bounce rate, conversion behaviour), which are critical for detecting account takeover and bot activity (Zeng Q. et al. [13] show that behavioral features contribute ~50.9% of predictive power in the fitted hybrid model); $X^{\inf o}$ – encompasses system-level events (access frequencies, IP patterns, protocol anomalies). Sophisticated threats manifest simultaneously across all three: a single-domain architecture is structurally unable to detect cross-domain patterns [10].

This decomposition is not merely a technical convenience but reflects the underlying economic structure of the system. Transactional data capture direct financial flows, behavioral data describe user interaction patterns, and system-level data provide contextual information about the operational environment. By preserving this structure, the model maintains interpretability and allows for domain-specific analysis.

Let $\mathcal{D} \subset \mathbb{R}^m$ de note the normal-behavior region estimated from training data. Anomaly detection is defined as a mapping from the feature space to a bounded interval, producing an anomaly score s_i for each observation x_i

$$f: \mathbb{R}^m \to [0,1],\ s_i = f(x_i) = P(x_i \notin \mathcal{D}) \quad (3)$$

In contrast to purely statistical interpretations, the anomaly score in this framework is treated as a proxy for the probability of abnormal economic behavior under incomplete information. This interpretation is crucial, as it establishes a direct link between statistical detection outputs and economic decision-making.

The estimation of anomaly scores is inherently uncertain, as the true distribution of normal behavior is unknown and may evolve over time. This uncertainty is particularly pronounced in digital business environments, where behavioral patterns are dynamic and influenced by external factors. As a result, anomaly detection must be understood as an approximation problem in which the function mapping observations to anomaly scores is learned from data rather than specified analytically.

To improve robustness, anomaly detection is performed separately within each data domain. This results in a set of domain-specific anomaly scores, each capturing different aspects of system behavior. The use of domain-level detection allows the model to account for heterogeneity in data distributions and to avoid the loss of information that may occur when features are aggregated prematurely. At the same time, it enables the identification of domain-specific anomalies that may have distinct economic interpretations.

The continuous formulation enables risk-stratified responses: observations with intermediate scores are queued for review rather than categorically blocked. The binary classification rule is:

$$\hat{y}_i = 1 if\ s_i > \tau;\ \hat{y}_i = 0 if\ s_i \le \tau \quad (4)$$

In the proposed framework, τ is an economic decision variable optimized against expected financial loss, not a statistical hyperparameter.

The economic interpretation of the anomaly score follows:

$$E[Loss \mid s_i] = s_i \cdot P(Threat \mid x_i) \cdot Impact(x_i), \quad (5)$$

where $s_i = f(x_i)$ is the domain-level anomaly score for observation x_i, representing the estimated probability that the observation deviates from the learned normal-behavior $\mathcal{D}$; P $P(Threat \mid x_i)$ is the conditional probability that the detected anomaly corresponds to an actual security threat, which may be estimated from historical incident data or expert elicitation; $Impact(x_i)$, is the expected financial exposure associated with observation x_i, reflecting transaction value, potential chargeback costs, and downstream operational losses.

The product $si \cdot P(Threat \mid x_i)$ yields the risk-adjusted anomaly likelihood, which is subsequently scaled by financial exposure to produce an economically interpretable detection signal. This formulation operationalizes the principle established by Gordon L. and Loeb M [3], that security decisions should be governed by expected loss rather than technical performance, at the level of individual observations.

User behaviour requires a profile-based deviation metric rather than a point-in-time feature comparison. For each user *u*, the temporal action sequence in window *t* is:

$$B_u(t) = \{a_1, a_2, \ldots, a_k\}, \quad (6)$$

where a_j denotes an individual user action (e.g., login, payment attempt, password reset, device change, or navigation event), *k* – is the number of recorded actions within the observation window, and *t* defines the analyzed time interval.

A normal profile B_u^{norm} is estimated from historical data across temporal, spatial, and functional dimensions. Deviation from this profile is:

$$D_u(t) = d\left(B_u(t), B_u^{norm}\right), \quad (7)$$

where $d(\cdot,\cdot)$ is cosine distance or KL-divergence for continuous vectors, and Levenshtein edit distance for sequential action patterns, which are normalized to [0, 1] to obtain the behavioral anomaly indicator $f^{beh}(x_i)$. This component detects account takeover, bot activity, and coordinated fraud preparation that may remain invisible at the transactional level [16].

The *domain-specific scores* (*domain level*) are subsequently combined into a unified multi-source indicator. The unified anomaly score is a weighted linear combination of domain-specific signals:

$$s_i^{MS} = \alpha \cdot f^{trans}(x_i) + \beta \cdot f^{\inf o}(x_i) + \gamma \cdot f^{\inf o}(x_i),\ \alpha + \beta + \gamma = 1, \quad (8)$$

where s_i^{MS} is the aggregated multi-source anomaly score for observation *i*; $f^{trans}, f^{beh}, f^{\inf o}$ denote anomaly scores obtained from transactional, behavioral, and system-level data domains, respectively; α, β, γ are non-negative weights reflecting the relative importance of each domain in the overall risk assessment; and $i = 1,2,\ldots,n$ indexes observations in the dataset.

Domain weights are estimated by minimising total expected economic loss on the validation set:

$$\left\{\alpha^*, \beta^*, \gamma^*\right\} = \arg\min Cost(\alpha, \beta, \gamma). \quad (9)$$

The use of a weighted linear aggregation is motivated by both analytical tractability and interpretability. From an economic perspective, the weights can be interpreted as reflecting the relative importance of different data domains in assessing risk. Their calibration can therefore be aligned with the objective of minimizing expected loss.

Cross-domain integration is necessary because sophisticated threats generate correlated anomalies across all three layers: thus, authors of [10; 11] identify multi-modal integration as the field's most impactful open gap.

Each domain function $f^{trans}, f^{beh}, f^{\inf o}$ is implemented through an ensemble of three algorithms with complementary detection mechanisms. Authors of [6] demonstrate empirically that heterogeneous ensembles outperform homogeneous ones because algorithms from different structural families exhibit complementary blind spots.

A fraud pattern engineered to evade one mechanism will, with high probability, be visible to at least one other, providing defence-in-depth in adversarial environments. Table 1 summarises the three component algorithms.

The anomaly detection component of the proposed framework relies on a combination of complementary unsupervised models, each capturing different aspects of abnormal behavior. Isolation Forest [17] identifies anomalies based on the principle of isolation, where observations that can be separated from the rest of the data with fewer partitioning steps are considered anomalous.

In contrast, the Autoencoder approach detects anomalies through reconstruction error, learning a compressed representation of normal behavior and flagging observations that cannot be accurately reconstructed as deviations from the underlying distribution.

Additionally, the One-Class SVM [18] defines a boundary around normal data in a transformed feature space and identifies observations lying outside this boundary as anomalies. The combination of these methods enables the framework to capture diverse anomaly patterns, improving robustness and detection performance across heterogeneous data environments.The three scores are aggregated into the *ML-level ensemble score*:

$$S_i^{ML} = w \cdot S_i^{IF} + w \cdot S_i^{AE} + w \cdot S_i^{SVM}, \ \Sigma w_j = 1, \quad (10)$$

where S^{ML} is the aggregated anomaly score for observation i; $S_i^{IF}, S_i^{AE}, S_i^{SVM}$ denote anomaly scores obtained from the Isolation Forest, Autoencoder, and One-Class SVM models, respectively; w_j are non-negative weights reflecting the relative importance of each model in the overall risk assessment; and $i = 1, 2, \ldots, n$ indexes.

The ensemble score is a convex combination of three algorithms outputs, with weights constrained to sum to one and optimized against the expected loss objective rather than statistical performance metrics. To complement the ensemble's sensitivity to novel anomalies, the framework incorporates a supervised component that estimates fraud probability from labeled historical observations.

The ensemble detects novel threats but does not exploit confirmed fraud labels when available. Let $\mathrm{P}(y_i = 1 \mid x_i)$ be the supervised fraud probability from a trained classifier. The hybrid score integrates both (*final detection level*):

$$S_i^{hybrid} = \lambda_1 S_i^{ML} + \lambda_2 P(y_i = 1 x_i), \lambda_1 + \lambda_2 = 1, \quad (11)$$

where S_i^{hybrid} is the unified anomaly score; S_i^{ML} is the ensemble-based anomaly score which is given by (10); $P(y_i = 1 x_i)$ is the supervised

Table 1

Complementary unsupervised algorithms in the detection ensemble

Algorithm	Mechanism	Strength	Limitation
Isolation Forest (IF)	Recursive binary isolation in feature space	Global point anomalies; high-dimensional data; O(n log n)	Insensitive to contextual/density anomalies
Autoencoder (AE)	Reconstruction error from compressed latent representation	Nonlinear deviations; complex behavioral patterns	Requires large normal-class training set
One-Class SVM (SVM)	Margin-based hypersphere in kernel space	Complex nonlinear decision surfaces	$O(n^2–n^3)$ scaling; subsampling required

Source: compiled by the authors based on [6; 17; 18; 20]

probability of fraud; and λ_1, λ_2 are non-negative weights controlling the relative contribution of unsupervised and supervised components. The weights are jointly optimized with the multi-source aggregation parameters α, β, γ respect to the economic loss function.

Thus, the hybrid score S_i^{hybrid} is the unified detection variable entering both the threshold rule and economic risk quantification. λ_1, λ_2 are jointly optimized with $\{\alpha, \beta, \gamma\}$ against the economic loss function. The supervised component detects known fraud patterns with high precision; the unsupervised ensemble detects novel threats regardless of historical labelling, together providing defence-in-depth against the full spectrum of e-commerce fraud typologies [11].

The proposed framework is operationalized as a multi-stage analytical pipeline, as illustrated in Figure 1. The system integrates heterogeneous data sources, including transactional, behavioral, and system-level information, which are jointly processed to capture different aspects of business activity.

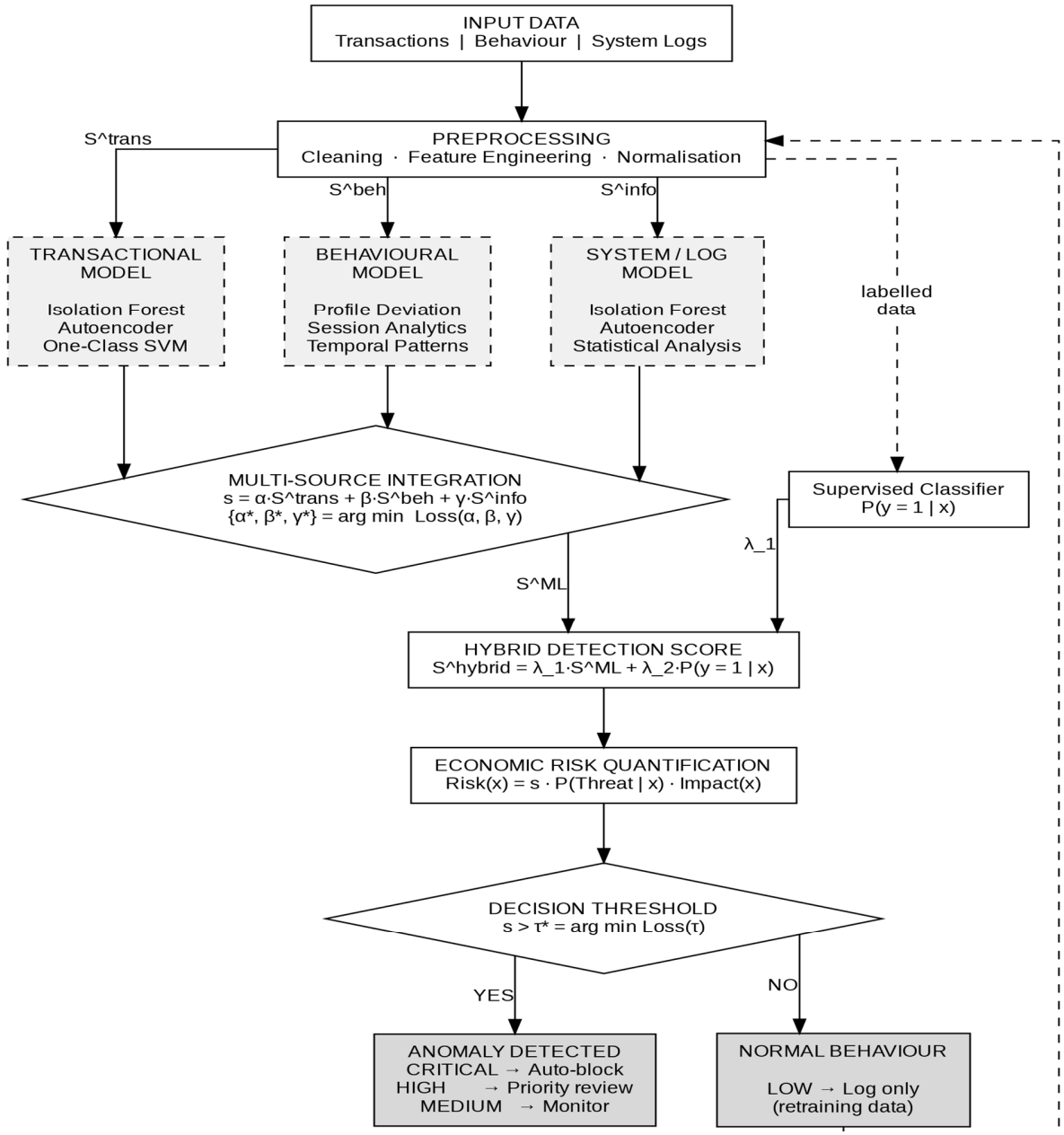

Figure 1. Multi-stage analytical pipeline for economic-aware anomaly detection: integrating heterogeneous data sources, hybrid detection, and cost-sensitive decision-making

Source: formed by the authors

At the *preprocessing stage*, data are cleaned, transformed, and normalized to ensure consistency across domains and to reduce the impact of noise and scale differences.

At the *detection stage*, anomaly scores are estimated independently within each data domain using complementary unsupervised models, enabling the capture of deviations across financial transactions, behavioral patterns, and system-level activities. The domain-specific scores are aggregated into a unified multi-source indicator with weights calibrated to minimize expected loss. A supervised classification component is incorporated to capture known fraud patterns, and its output is combined with the ensemble score into a hybrid anomaly score that balances sensitivity to novel threats with precision on known cases.

This hybrid score serves as the basis for economic risk quantification, combining anomaly likelihood with the estimated financial impact of each observation.

At the *final stage*, a cost-sensitive threshold determines the response category, from passive monitoring to active intervention, with adaptive retraining ensuring the model updates in response to evolving behavior.

The economic consequences of classification outcomes are inherently asymmetric: a missed fraudulent transaction incurs direct financial loss, chargeback costs, and operational expenses, whereas a false alarm primarily causes customer friction and reputational damage. Table 2 summarizes these outcomes based on representative industry estimates [2; 11; 19].

Empirical estimates indicate that the cost of a false negative significantly exceeds that of a false positive, with a typical ratio exceeding 7:1. This asymmetry implies that conventional evaluation metrics assigning equal weight to both error types may lead to economically suboptimal decisions. In particular, metrics such as the F1-score assigns equal importance to FP and FN errors, which is suboptimal when misclassification costs are asymmetric.

The total expected economic loss can be expressed as:

$$Loss = \sum_{i=1}^{n} \left[y_i \left(1 - \hat{y}_i\right) \cdot C_{FN} + \left(1 - y_i\right) \cdot \hat{y}_i \cdot C_{FP} \right]. \quad (12)$$

This formulation provides a direct link between classification outcomes and financial impact, enabling the calibration of detection models in terms of economic efficiency rather than statistical accuracy.

The corresponding optimization problem is defined as the minimization of expected loss with respect to both the detection function and the decision threshold:

$$\min_{f,\tau} Loss\left(f, \tau\right). \quad (13)$$

For a fixed detection model, the optimization reduces to selecting the decision threshold τ^* that minimizes expected loss:

$$\tau^* = \arg\min_{\tau} Loss\left(\tau\right). \quad (14)$$

A direct consequence of the cost-sensitive formulation is the repositioning of the classification threshold τ from a statistical hyperparameter to an economic decision variable. Under standard Bayes-risk minimization assumptions, the optimal threshold admits a closed-form solution (Höppner et al., 2022):

$$\tau^*_{analytical} = C_{FP} / \left(C_{FP} + C_{FN}\right) \approx 0.123. \quad (15)$$

This result implies that the optimal threshold is significantly lower than conventional values such as τ = 0.5, reflecting reflecting the 7:1 cost asymmetry: the system should flag a transaction as suspicious at a considerably lower evidence threshold than statistical convention would suggest, because the cost of missing a genuine threat far exceeds the cost of an unnecessary review. In practical deployment, τ^* is estimated empirically by minimizing the cost-weighted classification error on the validation set.

Table 2

Economic consequences of classification outcomes (e-commerce context)

Outcome	Condition	Economic consequence
True Positive (TP)	$y = 1, \hat{y} = 1$	Fraud blocked; avoided loss including transaction amount, chargeback, and operational costs
True Negative (TN)	$y = 0, \hat{y} = 0$	Legitimate transaction approved; no additional cost
False Negative (FN)	$y = 1, \hat{y} = 0$	Fraud not detected; financial loss including transaction value, chargeback, and operational costs
False Positive (FP)	$y = 0, \hat{y} = 1$	Legitimate transaction rejected; customer friction, support costs, and potential churn

Source: compiled by the authors based on [2; 11; 19]

The companion study demonstrates that thresholds of τ=0.3and τ=0.7produce economic losses 25% and 17% higher than the cost-sensitive optimum respectively, confirming that threshold calibration is not a secondary tuning decision but a primary source of economic value in deployed detection systems.ty.

In practical applications, the optimal threshold is estimated empirically using validation data by minimizing the cost-weighted sum of classification errors:

$$\tau^* = \arg\min_{\tau} \left[N_{FP(\tau)} \cdot C_{FP} + N_{FN(\tau)} \cdot C_{FN} \right]. \quad (16)$$

Empirical evidence shows that thresholds selected without accounting for cost asymmetry may lead to significantly higher economic losses. This highlights the importance of cost-sensitive calibration as a key component of anomaly detection systems in economic environments. This formulation directly links model calibration to economic decision-making criteria.

Binary classification alone is insufficient for operational decision-making in complex digital environments. To address this limitation, each observation is assigned an economic risk score that integrates detection results with financial impact:

$$Risk(x_i) = S_i^{hybrid} \cdot P(Threat|x_i) \cdot Impact(x_i), \quad (17)$$

where S_i^{hybrid} is the unified hybrid anomaly score defined in equation (11), integrating unsupervised ensemble evidence with supervised fraud probability; $P(Threat|x_i)$ retains the same interpretation as in equation (5); $Impact(x_i)$, denotes the financial exposure of observation x_i, consistently defined across both equations. The three components are treated as conditionally separable for modeling purposes, consistent with standard risk decomposition in security economics.

The overall system objective can therefore be expressed as the minimization of total expected risk:

$$Loss_{system} = \sum_{i=1}^{n} Risk(x_i) \rightarrow \min. \quad (18)$$

To operationalize this approach (Gordon L.A. and Loeb M.P. [3]), the continuous risk score is mapped to discrete decision levels that determine the appropriate response strategy. The corresponding decision logic is summarized in Table 3.

The distinction between equations (5) and (17) is architecturally significant: equation (5) operates at the domain detection stage, where s_i reflects a single-domain or pre-aggregation anomaly signal and serves to link raw detection outputs to economic interpretation; equation (17) operates at the decision stage, where S_i^{hybrid} incorporates the full multi-source aggregation and supervised integration pipeline, and serves as the basis for risk-tier assignment in Table 3.

This ensures that economic reasoning is embedded at two levels: as an interpretive principle governing model design, and as an operational metric governing response prioritization. This tiered decision structure allocates system resources efficiently, directing immediate intervention toward high-risk cases while lower-risk observations remain available for monitoring and model retraining.

Conclusions. The proposed framework extends the expected-loss principle from aggregate security investment to the internal design of detection systems. All architectural components, domain weighting, ensemble aggregation, hybrid integration, and threshold selection, are governed by a unified economic objective, shifting anomaly detection from a purely technical task to an economically grounded decision system. While cost-sensitive optimization has partially addressed this gap at the model level, the present approach extends it to the full detection and decision-making pipeline.

From a managerial perspective, the framework delivers three implications: multi-

Table 3

Risk-based decision matrix

Risk Level	Condition	Automated action	Human involvement
CRITICAL	$Risk > L_{crit}$	Automatic blocking; session termination	Immediate escalation; urgent review
HIGH	$L_{high} < Risk \le L_{crit}$	Transaction hold; additional verification	Priority review
MEDIUM	$L\_med < Risk \le L\_high$	Transaction allowed with flag	Deferred analysis
LOW	$Risk \le L_{med}$	No intervention	Logging and monitoring

Source: formed by the authors

source integration is essential as single-domain detection is structurally insufficient; the risk formulation translates technical outputs into financially interpretable indicators, improving communication between security teams and decision-makers; and economic threshold calibration aligns system behavior with organizational risk appetite.

Several limitations should be noted. Cost parameters are treated as population-level constants, whereas in practice they vary across transaction types and customer segments. Linear domain aggregation does not capture nonlinear cross-domain interactions. Empirical validation is currently limited to e-commerce and cybersecurity contexts and requires extension to other domains.

Overall, the framework reframes information security management as an economic governance problem driven by expected financial loss rather than statistical performance. Empirical evidence from the companion study confirms strong predictive and economic performance. Future research should address the implementation of the proposed approach into a multi-stage analytical pipeline that integrates data processing, anomaly detection, and decision-making into a unified system.

REFERENCES:

1. IBM Security. (2025). *Cost of a data breach report 2025*. IBM Corporation. https://www.ibm.com/reports/data-breach (accessed April 2, 2026).

2. Cybersource. (2024). *Global ecommerce payments & fraud report 2024*. https://www.cybersource.com/content/dam/documents/campaign/fraud-report/global-fraud-report-2024.pdf (accessed April 2, 2026).

3. Gordon, L. A., & Loeb, M. P. (2002). The economics of information security investment. *ACM Transactions on Information and System Security*, *5*(4), 438-457. https://doi.org/10.1145/581271.581274

4. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, *41*(3), 1-58. https://doi.org/10.1145/1541880.1541882

5. Aggarwal, C. C. (2017). *Outlier analysis* (2nd ed.). Springer. https://doi.org/10.1007/978-3-319-47578-3.

6. Goldstein, M., & Uchida, S. (2016). A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PLOS ONE*, *11*(4), Article e0152173. https://doi.org/10.1371/journal.pone.0152173.

7. Kumari, N., & Sami, A. (2024). A comprehensive investigation of anomaly detection methods in deep learning and machine learning: 2019-2023. *IET Information Security*, *2024*, Article 8821891. https://doi.org/10.1049/2024/8821891.

8. Baesens, B., Van Vlasselaer, V., & Verbeke, W. (2015). *Fraud analytics using descriptive, predictive, and social network techniques*. Wiley. https://doi.org/10.1002/9781119146841.

9. Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications*, *193*, Article 116429. https://doi.org/10.1016/j.eswa.2021.116429

10. Rodrigues, V. F., Becker, L. B., Bizotto, B. L., Canedo, E. D., Cardoso-Pereira, I., & de Mendonça, F. L. L. (2022). Fraud detection and prevention in e-commerce: A systematic literature review. *Electronic Commerce Research and Applications*, *56*, Article 101207. https://doi.org/10.1016/j.elerap.2022.101207.

11. Mutemi, A., & Bacao, F. (2024). E-commerce fraud detection based on machine learning techniques: Systematic literature review. *Big Data Mining and Analytics*, *7*(2), 419-444. https://doi.org/10.26599/BDMA.2023.9020023.

12. Byrapu Reddy, S., Jayaraman, R., Rao, B. D., & Prashanthi, J. (2024). Effective fraud detection in e-commerce: Leveraging machine learning and big data analytics. *Measurement: Sensors*, *33*, Article 101138. https://doi.org/10.1016/j.measen.2024.101138.

13. Zeng, Q., Lin, L., Jiang, R., Huang, W., & Lin, D. (2025). NNEnsLeG: A novel approach for e-commerce payment fraud detection using ensemble learning and neural networks. *Information Processing & Management*, *62*(1), Article 103916. https://doi.org/10.1016/j.ipm.2024.103916.

14. Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials*, *18*(2), 1153-1176. https://doi.org/10.1109/COMST.2015.2494502.

15. Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, *60*, 19-31. https://doi.org/10.1016/j.jnca.2015.11.016.

16. Maci, F., Coscia, P., Nicolardi, V., Ranieri, A., Rota, P., Sona, D., & Farinelli, A. (2024). A comprehensive investigation of clustering algorithms for UEBA. *Frontiers in Big Data*, *7*, Article 1375818. https://doi.org/10.3389/fdata.2024.1375818.

17. Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. *Proceedings of the 8th IEEE International Conference on Data Mining*, 413-422. https://doi.org/10.1109/ICDM.2008.17.

18. Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation*, *13*(7), 1443-1471. https://doi.org/10.1162/089976601750264965.

19. Höppner, S., Stripling, E., Baesens, B., vanden Broucke, S., & Verdonck, T. (2022). Instance-dependent cost-sensitive learning for detecting transfer fraud. *European Journal of Operational Research*, *297*(1), 291-300. https://doi.org/10.1016/j.ejor.2021.05.028.

20. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. https://www.deeplearningbook.org (accessed April 8, 2026).

Дата надходження статті: 21.04.2026
Дата прийняття статті: 12.05.2026
Дата публікації статті: 25.05.2026