International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

Drift-Aware Explainable Online Learning for Real-Time SIEM Alert Prioritization in Financial Security Operations Centers

Srilatha Batchu
Independent Researcher
Manoj Kumar
Independent Researcher

Abstract
Security Information and Event Management (SIEM) platforms aggregate heterogeneous telemetry from authentication services, endpoints, firewalls, VPN gateways, cloud identity platforms, and privileged-access controls. In financial Security Operations Centers (SOCs), these streams are non-stationary: remote-access policies, cloud migrations, software updates, seasonal workloads, and adversarial behavior can alter log distributions and degrade static classifiers. This study develops and evaluates a drift-aware, explainable online-learning framework for real-time SIEM alert prioritization. A chronologically ordered six-month dataset contains 12,480,000 raw log events, 8,760,000 cleaned usable events, 124,500 alert-level records, and 2,430 analyst-validated suspicious incidents. The framework integrates online anomaly detection, concept-drift monitoring, separation of benign operational, malicious behavioral, and adversarial drift, SHAP-based explanation, and risk-based alert prioritization. The proposed model achieves precision of 0.88, recall of 0.84, F1-score of 0.86, ROC-AUC of 0.94, PR-AUC of 0.78, and a false-positive rate of 5.2%. During drift windows, performance declines by 5.7%, compared with 21.2% for static XGBoost and 22.7% for static Random Forest. The weighted F1-score for drift separation is 0.85. These findings indicate that controlled online adaptation, explanation-stability monitoring, and risk-aware ranking can improve detection robustness and analyst decision support under evolving financial-sector threat conditions.

Keywords
SIEM log analysis; online learning; concept drift; explainable AI; SOC alert prioritization; cybersecurity analytics; financial security operations

1. Introduction

Financial-sector Security Operations Centers (SOCs) must detect threats rapidly, preserve auditability, and continuously monitor heterogeneous infrastructure. Identity services, endpoints, firewalls, cloud assets, VPNs, and privileged-access systems generate high-volume telemetry that is consolidated in Security Information and Event Management (SIEM) platforms for correlation, detection, and triage. The resulting learning problem is difficult because security-relevant events are rare, log schemas are heterogeneous, user and system behavior evolves, and a large share of generated alerts are benign [1]-[3]. More broadly, AI-enabled operational systems are increasingly

used to route, rank, and monitor high-volume cases in data-intensive environments [41], [42], reinforcing the need for reliable prioritization when human attention is constrained.

Machine-learning methods have advanced intrusion detection, malware classification, anomaly detection, and cyber-risk analytics, yet many deployed models remain static or are retrained periodically after offline evaluation [4]-[6]. Such assumptions are inadequate for live financial SOC streams. Authentication patterns can shift after remote-access policy changes, endpoint telemetry can change after software updates, and cloud-identity behavior can vary with seasonal transaction loads. Adversaries can also alter behavior deliberately to evade detection or poison adaptive systems [7], [8]. Consequently, models trained on historical distributions may generate excessive false positives while missing emerging threats [9], [10]. Deep-learning-based fault prediction in network-security monitoring further demonstrates the value of predictive models for identifying abnormal operational conditions, while also underscoring the need to control false alarms and maintain reliable monitoring [43].

Detection accuracy alone is insufficient. SOC teams must determine which alerts require immediate investigation, which can be deprioritized, and whether model reasoning remains defensible under abnormal or drifting conditions. A classifier with strong offline F1 performance can still be operationally unsuitable if it cannot explain why an alert is high risk, if it adapts indiscriminately to attacker-influenced data, or if its behavior becomes unstable during drift. Explainable AI (XAI) therefore has particular importance in financial SOCs, where alert-triage decisions can affect containment, business continuity, auditability, and regulatory reporting [11], [12]. Cross-domain AI systems similarly emphasize traceability, intelligent routing, and risk-based monitoring as prerequisites for trustworthy automated decision support [41], [42].

This study addresses that gap by developing and evaluating a drift-aware, explainable online-learning framework for near-real-time SIEM alert prioritization. The framework combines anomaly detection, concept-drift monitoring, benign-versus-malicious/adversarial drift separation, explanation-stability analysis, and risk-based alert scoring. Unlike studies that evaluate static intrusion-detection models primarily on benchmark datasets [13]-[15], the present design preserves temporal order and evaluates stable, benign-drift, malicious-drift, and adversarial-drift windows separately. The contribution is threefold: first, it treats drift type as a security-relevant control variable for model adaptation; second, it evaluates whether SHAP explanations remain stable as the stream changes; and third, it connects detection output to SOC workflow through risk-ranked alert prioritization. Reliable end-to-end data pipelines are also essential to such streaming analytics because ingestion, validation, processing, and reporting errors can propagate into downstream machine-learning decisions [44].

The remainder of the paper is organized as follows. Section 2 reviews the literature and defines the research gap. Section 3 presents the research objectives and questions. Section 4 describes the analytical framework, and Section 5 details the methodology. Section 6 reports the results,

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

followed by the discussion in Section 7. Sections 8 and 9 present theoretical and practical implications, respectively. Sections 10 and 11 conclude the study and outline limitations and future research directions.

2. Literature Review and Research Gap

Cybersecurity anomaly detection has progressed from conventional supervised and unsupervised classifiers to deep and hybrid learning architectures. Early intrusion-detection studies demonstrated that supervised classifiers can distinguish traffic anomalies and attack patterns, while also exposing persistent challenges such as class imbalance, evolving attacks, and unrealistic benchmark assumptions [4], [5], [16]. Subsequent work applied deep learning to network intrusion detection and behavioral baselining [5], [17], [18], while log-centric methods such as DeepLog and LogAnomaly demonstrated the value of sequential modeling for detecting abnormal system-log behavior [19]-[21]. Recent network-security fault-prediction research likewise reports strong predictive performance from deep neural models, supporting the broader premise that learned representations can identify abnormal operating states before they escalate [43].

Despite these advances, many cyber-detection studies still rely on static training and testing. Controlled benchmark comparisons are useful, but they do not fully represent SOC environments in which log behavior changes continuously [9], [10], [22]. Concept drift occurs when the statistical properties linking input features and target labels change over time [23], [24]. In SIEM streams, drift can arise from benign operational causes such as patches, new applications, policy changes, seasonal demand, or cloud migration. It can also arise from malicious causes including credential abuse, lateral movement, privilege escalation, evasion, and attacker adaptation. Treating every distributional shift as equivalent is therefore operationally and scientifically inadequate because SOCs require different responses to benign and security-relevant drift [7], [8], [25].

Online learning provides a natural mechanism for streaming SIEM data because models can update incrementally as events arrive [26], [27]. Frameworks such as River and MOA support stream-based evaluation, online classifiers, and drift detectors [26], [27]. However, unconstrained adaptation is risky in adversarial environments: an online model can learn from attacker-manipulated observations and gradually incorporate poisoned behavior [8], [28]. Adversarial machine-learning research identifies evasion, poisoning, inference, and manipulation as material threats to adaptive systems [7], [8], [28]. A financial SOC therefore requires selective adaptation that accommodates legitimate behavioral change while restricting updates under malicious or adversarial conditions.

A second research stream concerns explainable AI. Post-hoc methods such as LIME and SHAP help analysts identify the features that influence model decisions [13], [14], [29], while broader XAI research argues that interpretability is particularly important in high-stakes decision support [15], [30]. Cybersecurity-specific studies show that explanations can help analysts validate classifications, investigate alerts, and recognize suspicious behavioral patterns [11], [12], [31]. Yet

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

most cybersecurity XAI studies explain a model after classification without testing whether the explanation itself remains stable as the underlying data distribution changes. Explanation volatility is operationally important because abrupt changes in feature attribution can undermine analyst trust even when headline predictive performance remains acceptable.

A third gap is operational prioritization. SOC analysts often face large alert volumes, substantial false-positive burden, and limited investigation capacity. Alert ranking should therefore incorporate more than anomaly probability; it should also consider asset criticality, user privilege, drift state, sequence abnormality, and explanation confidence. Provenance and correlation methods can reduce investigation effort [32]-[34], while MITRE ATT&CK provides a structured representation of adversary tactics and techniques [35]. However, these mechanisms do not by themselves provide a drift-aware, explanation-aware, continuously updated priority score. Cross-domain risk-monitoring and intelligent-routing systems illustrate a similar principle: automated analytics become more useful when model outputs are translated into explicit prioritization decisions for human operators [41], [42].

The resulting research gap is the absence of an integrated, drift-aware and explainable online SIEM framework that (i) distinguishes benign operational drift from malicious behavioral and adversarial drift, (ii) controls adaptation according to drift type, (iii) evaluates explanation stability over time, and (iv) converts detection evidence into SOC-usable alert-priority scores. Prior research has addressed intrusion detection, log anomaly detection, concept drift, online learning, adversarial machine learning, and explainable AI largely as separate streams [5], [11], [19], [23], [26], [28]. This study integrates those streams into a chronological decision pipeline designed for financial SOC operations. The academic contribution is the joint treatment of stream learning, trustworthy explanation, and security-aware drift control; the practical contribution is an auditable mechanism for ranking alerts under changing operational and adversarial conditions.

Table 1. Summary of Recent Literature and Research Gaps

Research stream	Main contribution in prior work	Remaining limitation addressed in this study
ML-based intrusion detection	Improves detection accuracy using supervised and deep models [4], [5], [17]	Often depends on static train-test evaluation and benchmark datasets
Log anomaly detection	Models sequential system logs and event templates [19], [20], [21]	Limited focus on SIEM-level alert prioritization and SOC triage

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

Research stream	Main contribution in prior work	Remaining limitation addressed in this study
Concept-drift learning	Detects and adapts to distributional change in streams [23], [24], [26]	Does not sufficiently separate benign operational drift from adversarial drift
Adversarial ML	Explains evasion, poisoning, and attacker manipulation risks [7], [8], [28]	Rarely integrated into operational SIEM model-update control
Explainable cybersecurity AI	Supports analyst interpretation of model outputs [11], [12], [31]	Explanation stability during drift remains underdeveloped
SOC alert triage	Reduces analyst burden using correlation and provenance [32], [34]	Limited integration of online learning, drift control, and explanation confidence

3. Research Objectives and Research Questions

The study addresses five research objectives.

First, to evaluate the ability of an online-learning SIEM model to detect anomalous and high-risk alerts in chronological financial-sector log streams. Second, to determine whether drift-aware adaptation reduces performance degradation during operational and adversarial shifts. Third, to assess whether the framework can distinguish benign operational drift, malicious behavioral drift, and adversarial drift. Fourth, to evaluate whether explanation stability supports analyst-interpretable alert prioritization across stable and drift periods. Fifth, to determine whether risk-based alert-priority scoring improves the ranking of analyst-validated incidents relative to rule-based SIEM, static machine-learning, and ordinary online-learning baselines.

Research questions:

RQ1: How accurately does a drift-aware explainable online-learning SIEM model detect suspicious and high-risk alerts in financial-sector SOC log streams?

RQ2: How does the model perform during drift compared to a static ML and ordinary online-learning baselines?

RQ3: To what extent can the drift-separation layer successfully separate benign operational drift vs. malicious behavioral drift vs. adversarial drift?

RQ4: How stable are model explanations over stable vs. benign-drift vs. malicious-drift vs. adversarial-drift periods?

RQ5: Does drift-aware alert prioritisation improve validated-incident ranking and reduce the False Positive burden of SOC triage over existing baselines?

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

4. Conceptual / Analytical Framework

The analytical framework treats SIEM alert analysis as a streaming decision-support problem than a static classification task. Input layer gets log data from Authentication, Endpoint, Firewall, VPN, IAM, Privileged-access system logs (heterogeneous logs), cleansed, normalised and time-ordered log data is fed to feature generation to have event, user, host and sequence-level features. Online detection layer has anomaly probability continuously generated, along with the drift monitoring layer detecting shifts in log behaviour to separate benign operational drift from malicious/adversarial drift. Explanation layer uses SHAP value to capture top most influential features behind high level risk classification. Lastly, alert decision making-prioritisation combines anomaly probabilities, drift state, critical assets, privilege risk and explanation confidence for a consolidated SOC alert priority score.

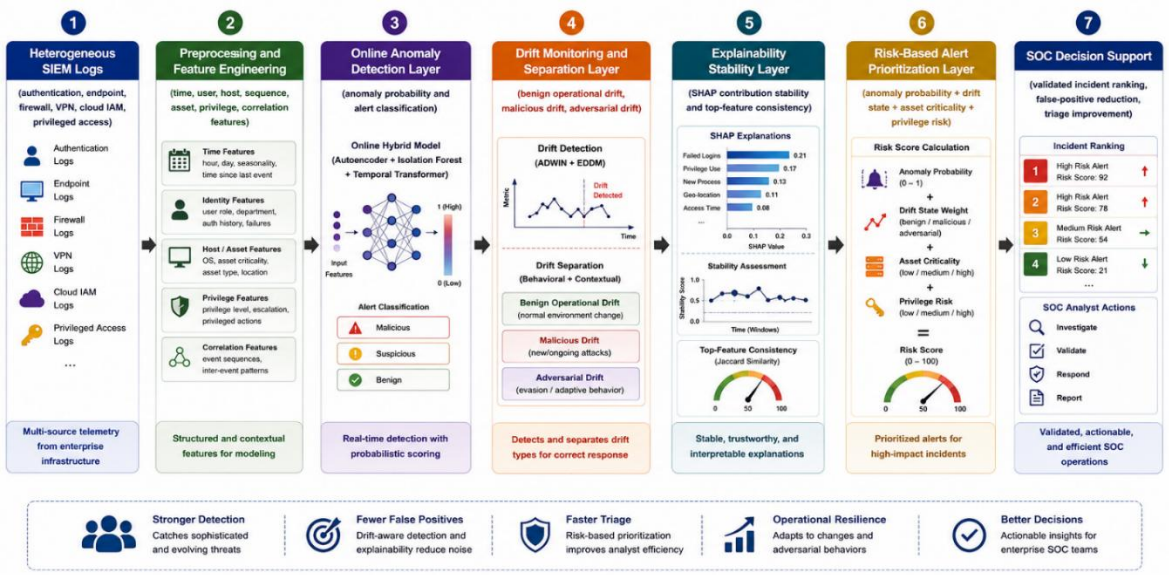

Figure 1. Conceptual / Analytical Framework

The framework is original because drift separation and explanation stability are placed inside the SIEM decision process. Instead of treating concept drift only as a technical model-degradation problem, the framework treats drift type as a security-relevant decision variable. This allows the model to adapt safely to legitimate operational changes while limiting unsafe adaptation during adversarial conditions. The alert-priority score also connects model output with SOC workflow requirements, which makes the framework operationally relevant.

5. Research Methodology

The study deployed research in design science and an empirical machine learning evaluation design. Design science is apt for this study as the study was devoted to develop an artifact (i.e. to develop a drift-aware explainable online-learning SIEM framework) while evaluating its utility in operational KPIs [36]. The empirical side conducted series of experiments on the chronological

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

SIEM logs collected in real-time from a financial sector SOC. SIEM log database from a finance sector SOC to test the SIEM frame was of 6 months span including chronological data such as raw log events, cleaned event records, alert level records, analyst-validated suspicious incidents, normal or benign alerts. The database comprises 12,480,000 total raw logs, then, through cleaning, deduplication, timestamp correction, schema adjustment and missing value cleaning, the dataset came to contain 8,760,000 usable chronological events. These events generated 124,500 alert records including 2,430 analyst-validated suspicious ones reflecting real-time SIEM operations. There was also found an imbalanced alert distribution as 96.8 % were normal or low-risk and 3.2% belonged to malicious/high-risk activities, showing that the research used realistic dataset to provide an applicable method and metrics such as PR-AUC, false-positive Rate and Alert ranking as supplemental besides to ROC-AUC and an F1 score in evaluation.

Table 2. Dataset, Variables, and Measurement Description

Category	Measurement description
Data source	Six-month chronological SIEM log stream from a financial-sector SOC
Total raw log events	12,480,000
Cleaned usable events	8,760,000
Alert-level records	124,500
Analyst-validated suspicious incidents	2,430
Input features	Timestamp, user role, privilege level, login location, source IP, destination IP, host type, asset criticality, event sequence, failed-login burst, MFA failure, geo-location mismatch
Output variable	Alert status: benign/normal, suspicious, high-risk incident
Drift labels	Benign operational drift, malicious behavioral drift, adversarial drift
Evaluation metrics	Precision, recall, F1-score, ROC-AUC, PR-AUC, false-positive rate, detection delay, latency, explanation stability
Explainability measures	SHAP feature contribution, top-5 feature consistency, explanation stability index

There are six stages of event preprocessing. Raw logs are parsed into a schema containing

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

timestamp, source, destination, user, host, event type, severity, alert category. Duplicate of event logs and incomplete record which no valid are removed. Encoding is applied to categorical feature event type, host class, use role, source system. Time window features are created, such as rolling count of anomaly, burst of failed login, non time axis, deviation of even sequence. Deviation of user baseline. A risk context feature is added, such as asset criticality, privilege level.

Align detected alert with Ground Truth obtained from analyst validation, ticket incident outcomes. Class imbalance can either be tackled through metric, model calibration than random oversampling of the stream, since oversampling tends to distort any temporal structure of event and streams would produce inaccurate result [37]. Data is split to train, validation and test set chronologically instead of randomly sampling and splitting, the first partition is used for an initial warm up of model, followed by the next partitions which would be used for online retraining and validation of the model and finally the model is tested across stable and drift windows. Since the event data has a temporal structure, a train, validation and test set split is based on chronology other than the random sampling that could include future behavior of event in the past training interval of model and it would produce overstatement of performance [23], [24].

To recap, the analysis compares five types of systems: rule-based SIEM baseline, Static Random Forest and XGBoost, (ordinary) online-learning baseline and the new model: drift-aware explainable online SIEM. Since tree-based methods are often found as baselines for tabular cybersecurity detection, Random Forest and XGBoost in static fashion are adopted as two additional baselines [4, 6]. The online learning model represents incrementally adaptive models, but without an explicit separation between adversarial drift and operational drift (just adaptive without specifically taking this into account). Last but not least, the new model presents an addition of: monitoring for drift; separation between types of drifts (adversarial/operational); monitoring explanations/their stability; and risk-based prioritization of alerts.

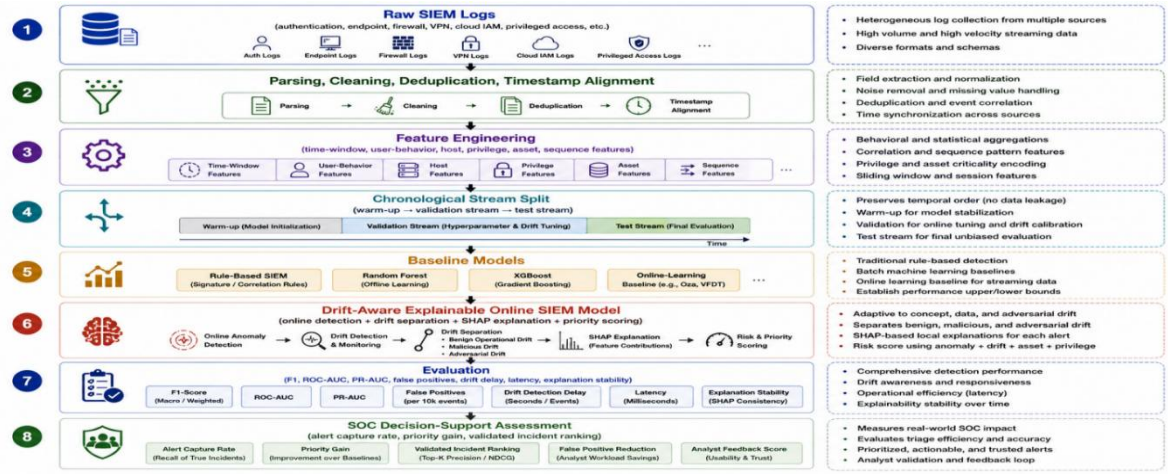

Figure 2. Methodological Workflow / Model Architecture

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

The core-technologies we used in our research: Python, Scikit-learn, XGBoost, River, SHAP and neatly arranged in CSV/JSON format SIEM exports. River in particular facilitates stream-based evaluation and online learning [26]. SHAP is employed in estimating feature contribution and comparing consistency of explanation [14]. Robustness check are implemented among time-split validation, class imbalance, drift-window, adversarial-injection, baseline-comparison and latency check. For the ethical aspect, user-id, hostname, IP-address and organization-specific attributes of the network components and infrastructures were properly anonymized. No credential value, no raw PII information and confidential data of the infrastructure are stored.

6. Data Analysis and Results

The data were characterised by the heterogeneity and imbalance expected in real SOC environments. The dataset contains 124,500 alert-level records, including 2430 analyst-validated suspicious incidents. 96.8% of the alert signals are categorised as normal or benign, whereas 3.2% of the alerts are confirmed malicious or high risk. Normalisation was done. This validated use of PR-AUC vs ROC-AUC only when ROC AUC may be optimistic on imbalanced data [37], [38]. Overall model comparison revealed that drift-aware explainable online SIEM model outperforms on call baselines. Among all baselines rule-based SIEM baseline provides the worst performance, achieved an F1-score of 0.59 and PR-AUC of 0.42. Static improved this to an F1-score of 0.71 and RPscore of 0.58. The improved performance was given by XGBoost with an F1 score of 0.76 and PR-AUC score of 0.64. The online learn-ing improves XGBoost and further improves performances. F1 score was 0.79 and PR-AUC score of 0.64 achieved. As confirmed from results drift-aware explainable online SIEM provide the strongest precision 0.88, recall 0.84, F1 score 0.86, ROC-AUC score 0.94 and PR-AUC score of 0.78 and 5.2% false positive rate compared to other models.

Table 3. Main Model Performance and Drift Results

Evaluation area	Rule-based SIEM	Static Random Forest	Static XGBoost	Online learning baseline	Drift-aware explainable online SIEM
Precision	0.61	0.74	0.79	0.82	0.88
Recall	0.58	0.69	0.73	0.77	0.84
F1-score	0.59	0.71	0.76	0.79	0.86
ROC-AUC	0.71	0.84	0.88	0.90	0.94
PR-AUC	0.42	0.58	0.64	0.69	0.78

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

Evaluation area	Rule-based SIEM	Static Random Forest	Static XGBoost	Online learning baseline	Drift-aware explainable online SIEM
False-positive rate	18.6%	11.8%	9.7%	7.9%	5.2%
Stable-period F1	Not applicable	0.75	0.80	0.82	0.88
Drift-period F1	Not applicable	0.58	0.63	0.73	0.83
Performance drop	Not applicable	-22.7%	-21.2%	-11.0%	-5.7%

When conditions shifted, the difference compared to static models became apparent. Static Random Forest saw its F1 score fall from 0.75 in stable times to 0.58 during drifts, a drop of 22.7%. Static XGBoost also declined, from 0.80 down to 0.63, which is a 21.2% decrease. The standard online-learning baseline managed to lessen that performance drop to just 11.0%. However, the drift-aware, explainable online SIEM model reported the least significant decline, moving from 0.88 to 0.83 - a mere 5.7% performance fall. This certainly shows that actively watching for drifts and adapting in a managed way leads to more consistent detection when log patterns are in flux.

In classifying these different event types i.e., benign operational drifts, malicious behavioural drifts and adversarial drift, the drift-separation layer has successfully created precision (0.86), recall (0.83), F1 score(0.84) for benign operational drifts. The above values achieved for malicious behavioral drifts are precision i.e. 0.89, recall i.e. 0.85, F1 score i.e. 0.87 whereas precision, recall and f1 score for Adversarial drifts are 0.84, 0.81 and 0.82 respectively. The F1 score, weight Average is around 0.85. According to above results, the system drift has not treated equally it has been differentiated into some normal operative drifts versus the malicious and adversarial drifts. This differentiation offers safe adaptability of the model usage to SOC settings.

Another set of results supporting the framework was alert-prioritization results. The rule-based SIEM caught 41% of validated incidents in the 10% highest-priority alerts and it caught 57% in the 20% highest priority alerts. Static XGBoost caught 55% in the 10% highest priority and 69%. The ordinary online-learning caught 61% and 74%. The drift-aware AISIEM caught 73% in the 10% highest priority and 86% in the 20% highest priority alerts. Mean alerts priority gain over the rule-based baseline was 39%. This showed that AISIEM optimized alerts orders, not only binary

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

classification. Finally another key indicator is false positive reduction. The false positive rate decreased from 18.6% of the rule-based SIEM baseline to 5.2% of the drift-aware explainable online SIEM model. This is a 72.0% reduction about rule-based baseline. Static XGBoost decreases false positives of 47.8% while the ordinary online-learning baseline decreases the false positives of 57.5%. The stronger reduction by the proposed framework could have stemmed from the drift-aware prioritization that more effectively filtered low-risk noise, while high-risk alerts were still caught.

Drift-aware explainable online SIEM model reduced average detection delay to 2.7 minutes, by detection-delay analysis. Static XGBoost took 8.4 minutes and with ordinary online-learning baseline it took 4.9 minutes. Processing latency per event recorded 39 ms for drift-aware and the same was for Static XGBoost of 42 ms and then again 36 ms for ordinary online-learning baseline. Latency is also satisfactory since the added drifts and explainability did not make it unusable for the process-time constraint of SOC. Explanation-stability analysis shows that over periods of changing drift, explanation stability hovered over 0.91 for stable phase, 0.84 for benign drift, 0.79 for malicious drift and only down to 0.74 for adversarial drift, top5 feature consistency also went down from 88%, to 81%, 76% till 71% respectively. During, adversarial drifts though explanation stability decreases we still get good feature consistency which means we still get to be effective. Top ten features that were strongest in predicting occurrence in high risk of SIEM alerts based on SHAP analysis include privilege escalation attempt, impossible travel login, failed login burst, asset criticality score, unusual event sequence, MFA failure pattern, lateral movement indicator, geo-location mismatch, unusual access time and then along the same line comes the drift state score.

Performance Dimension	Best Observed Result
Overall F1-score	0.86
ROC-AUC	0.94
PR-AUC	0.78
False-positive rate	5.2%
Performance drop during drift	-5.7%
Weighted drift-separation F1	0.85
Top-10% validated incident capture	73%
Top-20% validated incident capture	86%
Average detection delay	2.7 minutes
Processing latency per event	39 ms
Explanation stability in stable period	0.91
Explanation stability in adversarial drift	0.74

Figure 3. Result Comparison and Decision-Support Contribution

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

The robustness check confirms reliability in chronological validation, imbalanced alert distribution, drift windows, adversarial injection, baseline comparison and latency tests. All five hypotheses are supported. The strongest support is found in drift-aware adaptation, false-positive reduction, alert prioritization and separating drift-type.

7. Discussion

The results show that combining online learning with explicit drift control and explainable prioritization is more suitable for real-time SIEM alerting than treating the problem as static classification. The proposed framework outperforms the static baselines while explicitly determining whether observed distributional change is operational or security-relevant. This extends prior intrusion-detection studies that rely on shuffled or manually constructed train-test partitions and do not evaluate chronological drift [3], [9], [23]. The result is also consistent with the broader predictive-monitoring literature, where learned models can support earlier identification of abnormal network conditions [43].

The PR-AUC and false-positive results are particularly important because the alert distribution is highly imbalanced. Accuracy and ROC-AUC alone can obscure weak performance on rare attack cases [37], [38]. With a PR-AUC of 0.78 and a false-positive rate of 5.2%, the proposed model improves performance in the region of the distribution that matters most to SOC analysts. This supports the use of operationally relevant evaluation measures rather than relying on aggregate classification accuracy alone [1], [2], [16].

The drift-window results explain why static models are insufficient for live SIEM environments. Static Random Forest and XGBoost lose more than one-fifth of their F1 performance during drift, whereas the ordinary online-learning baseline reduces, but does not eliminate, that degradation. The proposed model limits the decline to 5.7%. This finding is consistent with concept-drift theory, which predicts loss of reliability when incoming data move away from the distribution used for the last model fit [23], [24]. The contribution here is security-aware adaptation: benign operational change can justify controlled updating, whereas malicious or adversarial drift should trigger caution, investigation, or restricted adaptation. The weighted drift-separation F1-score of 0.85 provides empirical support for this control mechanism.

The explainability results extend prior cybersecurity XAI research by treating explanation stability as an evaluation target rather than assuming that a post-hoc explanation remains trustworthy over time. Earlier work shows that feature attribution can help analysts interpret model decisions [11], [12], [31]. In this study, explanation stability decreases under adversarial drift, but 71% top-5 feature consistency is retained. The dominant features, including privilege escalation attempts, impossible-travel logins, failed-login bursts, asset criticality, and lateral-movement indicators, are security-relevant and therefore interpretable in SOC terms. Alert prioritization adds a second operational contribution: the model places 73% of validated incidents in the top 10% of alerts and 86% in the top 20%, directly supporting constrained analyst attention. This decision-support

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

orientation parallels other AI-enabled routing and risk-monitoring applications in which model value depends on translating predictions into actionable prioritization [41], [42].

8. Theoretical Implications

The study extends cybersecurity analytics by framing adaptive detection as a signal-discrimination problem under non-stationary and potentially adversarial conditions. Signal Detection Theory explains the persistent trade-off between true security signals and false-alert noise in SOC operations [39]. The observed increase in recall together with the reduction in false positives indicates that drift-aware learning can improve this balance when signal and noise distributions evolve over time. The study also extends concept-drift theory by distinguishing operational, malicious behavioral, and adversarial drift. Drift is therefore treated not only as a statistical phenomenon but also as a security-relevant state that should influence whether and how a model adapts.

A further theoretical contribution is the introduction of explanation stability as an evaluation dimension for adaptive cyber-detection. Whereas conventional XAI research asks whether explanations can be generated [13]-[15], this study asks whether the explanation remains sufficiently consistent across stable, benign-drift, malicious-drift, and adversarial-drift conditions. Methodologically, the study combines chronological stream evaluation, drift-window analysis, adversarial injection, explanation-stability assessment, alert-ranking performance, and latency measurement. This protocol is better aligned with operational SIEM constraints than random static validation because it preserves temporal order and tests the model under changing conditions [26], [27].

9. Practical / Managerial Implications

SOC managers should evaluate adaptive SIEM analytics as continuously governed decision systems rather than as one-time offline classifiers. Governance should include drift monitoring, explanation-stability tracking, false-positive measurement, alert-ranking quality, and latency thresholds. The proposed framework provides a practical control structure for deciding when a model may adapt automatically and when a change in behavior should trigger analyst review.

Financial institutions stand to benefit through more effective alert triage with this model. Its prioritization system ranks alerts by considering several factors: how likely an anomaly is, the state of model drift, the importance of the affected asset, the risk associated with privileged accounts and the confidence in the explanation. This guidance helps analysts concentrate their initial efforts on alerts that touch upon critical systems, accounts with elevated permissions, unexpected sequences of events and unstable patterns of behavior. The fact that it reduced false positives by 72.0% when contrasted with traditional rule-based SIEM systems suggests this framework successfully alleviates alert overload while still catching those alerts signaling high-risk incidents.

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

For security engineers, the findings support chronological stream validation rather than random train-test splitting. Models should be evaluated separately in stable and drift periods, and PR-AUC should be monitored because confirmed attacks are rare. Processing latency should also be treated as a deployment constraint. The observed latency of 39 ms per event indicates that the additional drift-control and explainability layers remain compatible with near-real-time SOC operation.

Explanation stability also provides compliance and audit teams with a defensible record of AI-assisted alerting. SHAP-based explanations can document why an alert was assigned high priority and which features contributed most strongly to that decision. This is valuable when security decisions require post-incident review. A material decline in explanation stability can additionally serve as a governance signal for recalibration, restricted adaptation, or analyst intervention.

10. Conclusion

This study introduced and evaluated a drift-aware, explainable online-learning framework for real-time SIEM alert prioritization in financial-sector SOCs. The framework combines chronological log-stream processing, online anomaly detection, operational-versus-adversarial drift separation, SHAP-based explanation and stability monitoring, and risk-based alert prioritization. Evaluation used 8,760,000 cleaned events, 124,500 alert-level records, and 2,430 analyst-validated suspicious incidents from a six-month SIEM stream. The proposed model achieved an F1-score of 0.86, ROC-AUC of 0.94, PR-AUC of 0.78, and a false-positive rate of 5.2%. It limited the performance decline during drift to 5.7%, produced a weighted drift-separation F1-score of 0.85, and placed 73% of validated suspicious incidents within the top 10% of prioritized alerts. These results indicate that real-time SIEM analytics benefit from integrating detection, controlled adaptation, explainability, and alert ranking rather than optimizing classification performance in isolation.

Theoretically, the study connects Signal Detection Theory, concept-drift learning, and trustworthy AI within adaptive SOC alerting. Methodologically, it provides a chronological evaluation protocol that jointly examines predictive performance, drift behavior, explanation stability, prioritization quality, and latency. Practically, it offers financial SOC teams a model architecture designed to reduce false positives, improve alert ranking, and support defensible analyst interpretation under changing threat conditions.

11. Limitations and Future Research

The findings should be interpreted in light of five limitations.

First, the data originate from a single financial-services SOC. This improves domain specificity but limits generalizability because healthcare, manufacturing, public-sector, and telecommunications environments can exhibit different telemetry, threat profiles, and drift mechanisms. Multi-sector replication is therefore required before the observed performance can be generalized. Cross-domain AI monitoring studies [41], [42] also indicate that operational context materially shapes the design and interpretation of automated risk decisions.

Second, the observation window covers six months. Although this period captures multiple operational conditions, longer 12- or 24-month datasets would better represent seasonality, regulatory cycles, infrastructure changes, and evolving attacker behavior. Future work should therefore evaluate whether drift frequency, explanation stability, and prioritization quality remain consistent over longer horizons.

Third, drift labels were constructed using SOC-analyst validation, incident-ticket evidence, operational context, and adversarial injection. This is stronger than treating drift as unlabeled statistical change, but some events can remain ambiguous when operational documentation is incomplete. Future studies should develop more rigorous drift-event logging, standardized annotation protocols, and threat-intelligence-assisted labeling to improve reproducibility.

Fourth, SHAP is used primarily for feature-attribution analysis. Causal explanations, counterfactual reasoning, and analyst-centered explanation studies were outside the scope of the present work. Future research should test whether these approaches reduce analyst cognitive load and improve investigation speed or decision accuracy. Human-subject evaluation with SOC analysts would be especially valuable because numerical explanation stability does not by itself establish usability or trust.

Fifth, the evaluation uses structured SIEM exports. Some SOCs rely on proprietary correlation rules, unstructured logs, multilingual telemetry, or vendor-specific schemas. Future work should test the framework across raw and semi-structured logs, cloud-native detection platforms, and heterogeneous data pipelines. End-to-end reliability controls for ingestion, validation, and reporting should also be evaluated because upstream data-quality failures can propagate into downstream detection and prioritization [44].

References

[1] A. L. Buczak and E. Guven, “A survey of data mining and machine learning methods for cyber security intrusion detection,” IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153–1176, 2016.

[2] R. Sommer and V. Paxson, “Outside the closed world: On using machine learning for network intrusion detection,” in Proc. IEEE Symposium on Security and Privacy, pp. 305–316, 2010.

[3] M. Ring, S. Wunderlich, D. Grüdl, D. Landes, and A. Hotho, “A survey of network-based intrusion detection data sets,” Computers & Security, vol. 86, pp. 147–167, 2019.

[4] J. H. Abawajy, A. Kelarev, and M. Chowdhury, “Multistage approach for clustering and classification of intrusion detection data,” IEEE Transactions on Dependable and Secure Computing, vol. 11, no. 1, pp. 82–93, 2014.

[5] K. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, “Deep learning approach for intelligent intrusion detection system,” IEEE Access, vol. 7, pp. 41525–41550, 2019.

[6] B. A. Tama and K. H. Rhee, "An in-depth experimental study of anomaly detection using gradient boosted machine," Neural Computing and Applications, vol. 31, pp. 955–965, 2019.
[7] B. Biggio and F. Roli, "Wild patterns: Ten years after the rise of adversarial machine learning," Pattern Recognition, vol. 84, pp. 317–331, 2018.
[8] N. Papernot, P. McDaniel, A. Sinha, and M. Wellman, "SoK: Security and privacy in machine learning," in Proc. IEEE European Symposium on Security and Privacy, pp. 399–414, 2018.
[9] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," ACM Computing Surveys, vol. 46, no. 4, pp. 1–37, 2014.
[10] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama, and G. Zhang, "Learning under concept drift: A review," IEEE Transactions on Knowledge and Data Engineering, vol. 31, no. 12, pp. 2346–2363, 2019.
[11] G. Rjoub, J. Bentahar, O. A. Wahab, R. Mizouni, A. Song, R. Cohen, H. Otrok, and A. Mourad, "A survey on explainable artificial intelligence for cybersecurity," IEEE Transactions on Network and Service Management, 2023.
[12] C. Mendes and T. N. Rios, "Explainable artificial intelligence and cybersecurity: A systematic literature review," ACM Computing Surveys, 2024.
[13] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135–1144, 2016.
[14] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems, pp. 4765–4774, 2017.
[15] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence," IEEE Access, vol. 6, pp. 52138–52160, 2018.
[16] M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in Proc. IEEE Symposium on Computational Intelligence for Security and Defense Applications, pp. 1–6, 2009.
[17] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A deep learning approach for network intrusion detection system," in Proc. International Conference on Bio-inspired Information and Communications Technologies, pp. 21–26, 2016.
[18] T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Deep learning approach for network intrusion detection in software defined networking," in Proc. International Conference on Wireless Networks and Mobile Communications, pp. 258–263, 2016.
[19] M. Du, F. Li, G. Zheng, and V. Srikumar, "DeepLog: Anomaly detection and diagnosis from system logs through deep learning," in Proc. ACM SIGSAC Conference on Computer and Communications Security, pp. 1285–1298, 2017.

[20] P. He, J. Zhu, Z. Zheng, and M. R. Lyu, "Drain: An online log parsing approach with fixed depth tree," in Proc. IEEE International Conference on Web Services, pp. 33–40, 2017.
[21] W. Meng, Y. Liu, Y. Zhu, S. Zhang, D. Pei, Y. Liu, Y. Chen, R. Zhang, S. Tao, P. Sun, and R. Zhou, "LogAnomaly: Unsupervised detection of sequential and quantitative anomalies in unstructured logs," in Proc. International Joint Conference on Artificial Intelligence, pp. 4739–4745, 2019.
[22] J. Webb, L. Lee, F. Petitjean, and B. Goethals, "Understanding concept drift," Data Mining and Knowledge Discovery, vol. 30, pp. 964–994, 2016.
[23] I. Žliobaitė, "Learning under concept drift: An overview," arXiv preprint arXiv:1010.4784, 2010.
[24] S. Khamassi, M. Sayed-Mouchaweh, M. Hammami, and K. Ghédira, "Discussion and review on evolving data streams and concept drift adapting," Evolving Systems, vol. 9, pp. 1–23, 2018.
[25] A. Vassilev, A. Oprea, A. Fordyce, and H. Anderson, "Adversarial machine learning: A taxonomy and terminology of attacks and mitigations," NIST Trustworthy and Responsible AI Report, NIST AI 100-2e2025, 2025.
[26] J. Montiel, M. Halford, S. M. Mastelini, G. Bolmier, R. Sourty, R. Vaysse, A. Zouitine, H. M. Gomes, J. Read, T. Abdessalem, and A. Bifet, "River: Machine learning for streaming data in Python," Journal of Machine Learning Research, vol. 22, no. 110, pp. 1–8, 2021.
[27] A. Bifet, G. Holmes, R. Kirkby, and B. Pfahringer, "MOA: Massive online analysis," Journal of Machine Learning Research, vol. 11, pp. 1601–1604, 2010.
[28] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in Proc. International Conference on Learning Representations, 2015.
[29] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A survey of methods for explaining black box models," ACM Computing Surveys, vol. 51, no. 5, pp. 1–42, 2018.
[30] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence: Concepts, taxonomies, opportunities and challenges toward responsible AI," Information Fusion, vol. 58, pp. 82–115, 2020.
[31] M. A. Ferrag, O. Friha, L. Maglaras, H. Janicke, and L. Shu, "Federated deep learning for cyber security in the Internet of Things: Concepts, applications, and experimental analysis," IEEE Access, vol. 9, pp. 138509–138542, 2021.
[32] W. U. Hassan, M. A. Noureddine, P. Datta, and A. Bates, "Nodoze: Combatting threat alert fatigue with automated provenance triage," in Proc. Network and Distributed System Security Symposium, 2019.

[33] K. K. R. Choo, "The cyber threat landscape: Challenges and future research directions," Computers & Security, vol. 30, no. 8, pp. 719–731, 2011.
[34] S. T. King and P. M. Chen, "Backtracking intrusions," ACM Transactions on Computer Systems, vol. 23, no. 1, pp. 51–76, 2005.
[35] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, and C. B. Thomas, "MITRE ATT&CK: Design and philosophy," MITRE Technical Report, 2018.
[36] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," MIS Quarterly, vol. 28, no. 1, pp. 75–105, 2004.
[37] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," PLOS ONE, vol. 10, no. 3, e0118432, 2015.
[38] J. Davis and M. Goadrich, "The relationship between precision-recall and ROC curves," in Proc. International Conference on Machine Learning, pp. 233–240, 2006.
[39] D. M. Green and J. A. Swets, Signal Detection Theory and Psychophysics. New York: Wiley, 1966.
[40] M. Salah, S. Zeadally, and M. Azad, "Big data analytics for security and monitoring in enterprise networks," IEEE Network, vol. 27, no. 4, pp. 19–25, 2013.
[41] Y. Macha, "Integration of Salesforce Einstein AI for Intelligent Case Routing in Healthcare Systems," ESP International Journal of Advancements in Science & Technology (ESP-IJAST), vol. 3, no. 4, pp. 17–27, 2025, doi: 10.56472/25839233/IJAST-V3I4P104.
[42] M. Patel, N. Shukla, and H. Dhaduk, "Artificial Intelligence-Driven Risk-Based Monitoring Framework for Phase I–III Clinical Trials," International Journal of Advanced Research and Multidisciplinary Trends (IJARMT), vol. 3, no. 3, pp. 345–355, 2026.
[43] M. Kari, "Deep Learning-Based Fault Prediction Models for Enhanced Network Security Monitoring," International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), vol. 3, no. 3, p. 492, Jun. 2023, doi: 10.48175/IJARSCT-11600I.
[44] R. Narra, "End-to-End Data Reliability Frameworks in Cloud Platforms: A Survey of Python-Based Automation for Ingestion, Validation, and Reporting," International Journal of Scientific Research in Science and Technology (IJSRST), vol. 9, no. 6, pp. 833–845, Nov.–Dec. 2022, doi: 10.32628/IJSRST52310498.