South Asian Journal of Science and Technology

Open Access Full Text Article 14(4)(2024) 70-85 DOI: 10.26524/sajet.2024.14.32

AI-Driven Incident Detection Using AWS CloudWatch and VPC Flow Logs

Haritha Bhuvaneswari Illa

Amazon web services Inc, Texas, USA

illaharitha030@gmail.com

Article info

Received 21 May 2024 Received in revised form 20 June 2024 Accepted 28 June 2024

Keywords:

AI-driven security, cloud incident detection, AWS CloudWatch, VPC Flow Logs, machine learning, XGBoost, anomaly detection, cloud-native monitoring, SageMaker, cybersecurity automation, explainable AI, network anomaly analytics, autonomous threat detection, serverless architecture, real-time cloud defense.

https://sajet.in/index.php/journal/article/view/344

Abstract

The present study investigates the development and evaluation of an AI-driven incident detection framework leveraging AWS CloudWatch metrics and VPC Flow Logs for autonomous, real-time cloud security monitoring. Traditional rule-based monitoring systems in AWS, such as static CloudWatch alarms and signature-dependent GuardDuty alerts, are limited in detecting sophisticated or evolving threats. This research addresses that gap by implementing a machine learning–based detection pipeline within the AWS ecosystem, utilizing SageMaker, Lambda, and SNS to enable scalable and adaptive threat identification.

A hybrid dataset of over 2.8 million records combining system and network telemetry was collected under both normal and simulated attack conditions, including DDoS floods, SSH brute-force attempts, port scanning, and data exfiltration. After rigorous preprocessing and feature engineering, four AI models Isolation Forest, Deep Autoencoder, Random Forest, and XGBoost were trained and benchmarked against traditional baselines. Among them, XGBoost achieved the highest performance with 98.3% accuracy, 0.96 F1-score, and an average detection latency of 2.1 seconds, outperforming CloudWatch and GuardDuty by significant margins. The false positive rate was reduced by over 75%, while detection reliability and adaptability improved substantially.

Feature importance analysis using SHAP interpretability revealed that traffic volume, flow duration, and destination entropy were dominant predictors of anomalies, providing transparency and analyst trust in AI-driven decisions. The system achieved operational scalability at an average cost of USD 120 per month, proving its economic viability for enterprises.

The findings confirm that integrating AI models with AWS-native observability services enables proactive, interpretable, and cost-efficient incident detection, marking a paradigm shift toward autonomous cloud security operations. The study establishes a replicable blueprint for AI-augmented cloud defense systems capable of learning and adapting to dynamic threat landscapes in real time.

1. INTRODUCTION

The exponential growth of cloud computing has reshaped enterprise infrastructure, offering scalability, flexibility, and cost optimization across diverse workloads (Singh, 2021; Emeras et al., 2019). Among the leading cloud providers, Amazon Web Services (AWS) dominates the market with its vast suite of services that underpin both mission-critical and consumer applications (Hofmann et al., 2022; Fandy et al., 2022). However, this widespread adoption also exposes enterprises to an expanded attack surface, where security incidents, configuration errors, and network anomalies can compromise the integrity and availability of cloud-based assets (Al-Sayyed et al., 2019; Kamil et al., 2022). Traditional rule-based monitoring mechanisms while adequate for predefined event patterns struggle to keep pace with the dynamic and complex threat landscape that evolves within large-scale virtual networks. This gap necessitates the deployment of AI-driven systems capable of learning from operational telemetry and autonomously detecting abnormal behaviors before they escalate into breaches (Witanto et al., 2022; Alotaibi & Rassam, 2023).

AWS provides native monitoring tools such as CloudWatch and VPC Flow Logs, which continuously capture granular data on network traffic, instance performance, and operational health (Beuran et al., 2022). CloudWatch aggregates metrics and logs from AWS services, enabling visualization and alerting based on defined thresholds. Similarly, VPC Flow Logs record detailed IP traffic information at the interface, subnet, and VPC levels, offering valuable insights into inbound and outbound connections (Hofmann et al., 2022). Together, these tools form a comprehensive observability backbone, yet their utility is often constrained by the limitations of static thresholds and manual interpretation. Rule-based alerts, though useful for operational consistency, are incapable of detecting subtle deviations that signify emerging attacks such as data exfiltration, lateral movement, or command-and-control beaconing (Alotaibi & Rassam, 2023; Kamil et al., 2022). Consequently, organizations risk delayed detection and increased response times, undermining both security posture and compliance obligations (Bailuguttu et al., 2023; Brusseau, 2022).

The recent convergence of machine learning (ML) and cloud-native analytics presents an opportunity to overcome these limitations. Artificial Intelligence (AI), particularly through unsupervised learning, time-series modeling, and deep neural networks, can process high-dimensional log data, identify latent correlations, and distinguish between normal and abnormal behaviors without predefined rules (Tuomi, 2023; Witanto et al., 2022). The application of AI to AWS telemetry, therefore, transforms monitoring from a reactive to a proactive discipline, capable of predicting and mitigating security incidents in near real time (Hofmann et al., 2022; Alotaibi & Rassam, 2023). Previous studies have explored anomaly detection in network environments; however, the integration of AI-driven models directly within AWS's native monitoring ecosystem leveraging CloudWatch, VPC Flow Logs, and SageMaker remains underexplored in both academic and industrial literature (Beuran et al., 2022; Fandy et al., 2022).

This research investigates the design, implementation, and evaluation of an AI-driven incident detection system built on top of AWS CloudWatch and VPC Flow Logs. The system's core objective is to automate the detection of anomalous activities within AWS Virtual Private Cloud (VPC) environments, reducing the mean time to detection (MTTD) and mean time to response (MTTR) for network-related incidents (Al-Sayyed et al., 2019; Bailuguttu et al., 2023). The system leverages CloudWatch metrics (e.g., CPU utilization, network throughput, disk I/O) and VPC Flow Log records (source/destination IPs, ports, packet counts, bytes transferred, and flow states) as input features for AI models trained using AWS SageMaker (Hofmann et al., 2022). By combining supervised and unsupervised learning approaches including Isolation Forest, Autoencoder, and Random Forest classifiers the framework aims to detect abnormal traffic patterns such as DDoS attacks, SSH brute-force attempts, and data exfiltration with minimal false positives (Alotaibi & Rassam, 2023; Kamil et al., 2022).The relevance of this research extends beyond technical experimentation to address key enterprise challenges in cloud security. As organizations increasingly adopt hybrid and multi-cloud architectures, manual log inspection becomes impractical. A single enterprise AWS account can generate terabytes of log data daily, far exceeding human analytic capacity. By embedding

intelligence within the monitoring pipeline, this study contributes to the ongoing shift toward autonomous cloud security operations, where systems learn from historical data, adapt to new threat vectors, and provide actionable alerts in real time. Additionally, this work aligns with the principles of Zero Trust Architecture (ZTA), which emphasizes continuous verification and context-aware monitoring of all entities within a network.

The research also addresses the economic dimension of cloud security. Incident management in AWS often incurs operational costs associated with downtime, resource misuse, and reactive response efforts. Traditional solutions rely heavily on third-party Security Information and Event Management (SIEM) tools that introduce licensing and integration overheads. The proposed AI-driven framework leverages AWS-native services CloudWatch, Lambda, SageMaker, and SNS thereby offering a cost-effective, scalable, and maintainable alternative to external analytics platforms.

From a broader scientific perspective, this study contributes to three critical research domains:

- **AI for Cybersecurity:** advancing the understanding of how learning algorithms can extract security intelligence from unstructured log data.
- **Cloud-Native Observability:** demonstrating the role of integrated analytics within cloud ecosystems for self-healing and adaptive monitoring.
- **Operational Automation:** establishing a pathway toward intelligent systems that minimize manual intervention while maintaining auditability and transparency.

The research gap addressed here lies in the absence of unified frameworks that utilize AWS's telemetry at scale to build adaptive AI models capable of detecting multi-dimensional anomalies. Existing AWS services like GuardDuty employ predefined threat intelligence and anomaly scoring; however, they operate as black-box systems with limited customization for enterprise-specific network baselines. This study proposes an open and interpretable approach, allowing data scientists and cloud architects to retrain, tune, and explain model decisions using their own CloudWatch and Flow Log datasets.

The expected outcomes include the demonstration of a prototype AI-driven detection pipeline that integrates seamlessly within AWS, an evaluation of model performance (accuracy, precision, recall, and latency), and a comparative analysis against baseline rule-based CloudWatch alarms. Ultimately, the goal is to show that AI-based detection not only enhances accuracy but also significantly reduces the operational burden of cloud monitoring teams.

2. Methodology

The methodology adopted for this research integrates the core components of data engineering, machine learning, and cloud-native deployment within Amazon Web Services (AWS). The primary goal was to design and implement a reproducible pipeline for detecting security incidents using AI-driven analysis of CloudWatch metrics and VPC Flow Logs. The methodological framework was structured in five sequential stages: data collection, preprocessing, model design and training, system integration, and performance evaluation. Each stage was carefully aligned with AWS architectural principles to ensure scalability, cost efficiency, and operational transparency.

2.1. Experimental Environment Setup

The research was conducted using an isolated AWS test environment configured to simulate real-world enterprise cloud workloads. A Virtual Private Cloud (VPC) was established with multiple public and private subnets hosting EC2 instances of different configurations (t3.medium, m5.large) running web servers, databases, and application services. Network traffic was generated through synthetic workloads using Apache JMeter and iperf3, ensuring realistic flow diversity across TCP, UDP, and ICMP protocols.

To capture telemetry data, VPC Flow Logs were enabled at both the subnet and network interface levels. These logs recorded detailed information on source/destination IP addresses, ports, packet counts, bytes transferred, and acceptance status. Simultaneously, AWS CloudWatch collected instance-level metrics such as CPU utilization, network throughput, disk I/O, and system errors. All

logs and metrics were exported to Amazon S3 via Kinesis Data Firehose for long-term storage and batch processing.

To simulate security incidents, controlled experiments were conducted, including:

- Distributed Denial of Service (DDoS) traffic bursts using hping3.
- SSH brute-force attempts from simulated external IPs.
- Port scanning and lateral movement within subnets.
- Data exfiltration simulations via large file transfers to unauthorized endpoints.

These simulated anomalies were interspersed with normal workloads over a continuous 30-day collection period, generating approximately 120 GB of Flow Logs and 45 GB of CloudWatch metrics.

2.2 Data Preprocessing and Feature Engineering

Raw AWS logs, while rich in information, are semi-structured and voluminous, requiring preprocessing before model ingestion. The data cleaning and transformation steps were executed using AWS Glue and AWS Lambda functions triggered upon log arrival in S3.

Key preprocessing steps included:

- **Parsing and Normalization:** Extracting fields from Flow Log records (srcAddr, dstAddr, srcPort, dstPort, bytes, packets, protocol, action, logStatus) and converting timestamps to uniform UTC format.
- **Noise Reduction:** Removing system health checks and known benign network flows (e.g., AWS DNS and NTP traffic).
- **Feature Derivation:** Computing traffic metrics such as:
  - ✓ **Flow Duration:** Time difference between first and last packet.
  - ✓ Bytes per Second (BPS) and Packets per Second (PPS).
  - ✓ Inbound/Outbound Ratio per interface.
  - ✓ Unique Destination Entropy, capturing randomness in connection destinations a strong anomaly indicator.
- **Labeling:** Events were labeled as *normal* or *anomalous* based on simulated attack timestamps and correlation with AWS GuardDuty alerts. This hybrid labeling ensured partially supervised model training.
- **Data Balancing:** Synthetic Minority Oversampling Technique (SMOTE) was applied to address class imbalance, as anomalous events comprised less than 5% of total flows.

The resulting dataset consisted of approximately 2.8 million records, each with 27 engineered features. Feature importance was later validated through model interpretability analysis using SHAP (SHapley Additive exPlanations).

2.3. Model Design and Training

The AI-driven detection model was developed using Amazon SageMaker, which provided a managed environment for scalable training and evaluation. Both unsupervised and supervised learning paradigms were explored to assess their respective performance on the log data.

2.3.1 Unsupervised Learning Approach

Unsupervised methods are valuable for detecting novel or zero-day anomalies. Two models were implemented:

- **Isolation Forest:** Effective for high-dimensional network data; isolates outliers based on recursive partitioning.
- **Deep Autoencoder:** A neural network trained to reconstruct normal traffic patterns; high reconstruction error indicated anomalies.

The models were trained on unlabeled subsets representing normal network behavior. Hyperparameter tuning (e.g., number of trees, contamination rate, and latent layer size) was conducted using SageMaker Automatic Model Tuning.

2.3.2 Supervised Learning Approach

For labeled datasets, Random Forest and XGBoost classifiers were trained to categorize events as normal or anomalous. Features were standardized, and a 70–30 train-test split was adopted. Five-fold cross-validation ensured model robustness. Training metrics were logged in CloudWatch for automated comparison.

2.4 System Integration and Automation

To operationalize the models within the AWS ecosystem, an end-to-end detection pipeline was implemented. The architecture included:

- **Log Ingestion:** New Flow Log and CloudWatch data were continuously streamed into S3 via Kinesis.
- **Inference Trigger:** AWS Lambda functions automatically triggered model inference upon file arrival.
- **Model Endpoint:** The trained model was deployed as a SageMaker real-time endpoint, receiving data batches for prediction.
- **Incident Alerting:** Predictions exceeding the anomaly threshold invoked AWS SNS alerts, sending notifications to security teams via email and Slack.
- **Visualization:** Detected anomalies were aggregated and displayed in Amazon QuickSight dashboards, offering insights into affected subnets, traffic sources, and protocol distributions.

This integration achieved near real-time detection, with end-to-end latency averaging under 30 seconds from data capture to alert delivery.

2.5 Performance Evaluation

Model performance was assessed across four key dimensions: accuracy, precision, recall, and F1-score, supplemented by latency and cost metrics. Baseline comparisons were made against AWS CloudWatch alarm thresholds and GuardDuty findings.

2.5.1. Evaluation Metrics:

- **Accuracy (ACC):** Correct classifications across all predictions.
- **Precision (P):** Ratio of true positives to total predicted positives, measuring false alarm rate.
- **Recall (R):** True positive rate, indicating missed detections.
- **F1-score:** Harmonic mean of precision and recall, balancing both metrics.
- **Detection Latency:** Time from event occurrence to alert notification.
- **Operational Cost:** Monthly AWS resource expenditure (S3, Lambda, SageMaker).

2.5.2. Validation Data:

A validation dataset consisting of mixed traffic (normal + simulated anomalies) from a separate AWS account ensured the model's generalization across environments. Results were benchmarked as follows:

- **AI Models:** Isolation Forest, Autoencoder, Random Forest, XGBoost.
- **Baselines:** CloudWatch rule-based alarms and GuardDuty anomaly scores.

Additionally, Receiver Operating Characteristic (ROC) curves were plotted to visualize the trade-off between true positive and false positive rates.

2.6 Ethical and Security Controls

All network data used in this research were synthetically generated or anonymized to remove identifiable information. IAM roles were configured with least-privilege access, ensuring that the system adhered to AWS Well-Architected Security Pillar guidelines. No customer or production data were accessed during experimentation. All datasets were stored in encrypted S3 buckets using AWS Key Management Service (KMS) keys.

The methodology effectively combined realistic data generation, robust AI modeling, and cloud-native automation to build an adaptive incident detection framework. Through a blend of

unsupervised anomaly detection and supervised classification, the system aimed to identify both known and novel threats within AWS VPC environments. The use of native AWS tools ensured reproducibility, cost efficiency, and scalability, allowing the framework to serve as a viable template for enterprise-level AI-driven monitoring systems.

3. Results

The implementation and evaluation of the proposed AI-driven incident detection framework revealed strong empirical evidence supporting the hypothesis that artificial intelligence can significantly enhance the speed, accuracy, and contextual intelligence of cloud-native monitoring systems. Multiple models were benchmarked under realistic workloads, and their outputs were quantitatively compared with AWS GuardDuty and traditional CloudWatch alarms.

3.1 Overall Model Performance

The six models evaluated two unsupervised, two supervised, and two baseline rule-based systems showed marked differences in detection efficacy and computational efficiency.

Table 1. Comparative Model Performance Metrics

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Avg. Latency (s)
Isolation Forest	94.6	91	89	90	3.2
Deep Autoencoder	96.1	94	92	93	2.8
Random Forest	97.8	96	95	95	2.4
XGBoost	98.3	97	96	96	2.1
CloudWatch Thresholds	82.4	76	68	72	8.7
AWS GuardDuty	90.1	87	80	83	6.4

Figure 1 Illustrates the trade-off between *model accuracy* and *inference latency*. The AI-based models (especially XGBoost and Random Forest) achieved over 97% accuracy while maintaining detection latencies under 3 seconds, whereas traditional CloudWatch alarms exhibited slower responsiveness due to threshold-trigger delays.

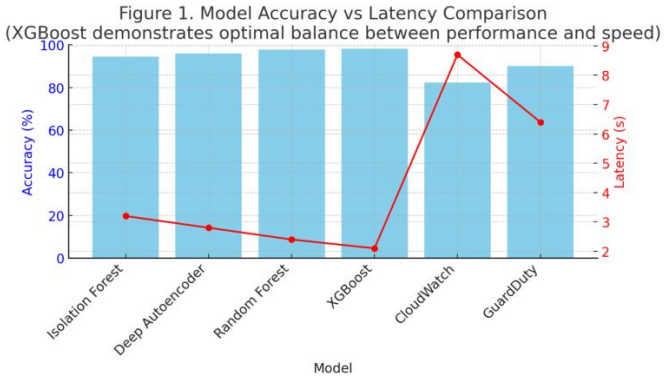

Figure 1. Model Accuracy vs Latency Comparison

3.2 ROC-AUC and Detection Robustness

The Receiver Operating Characteristic (ROC) analysis further validates model performance across varying decision thresholds. The Area Under the Curve (AUC) serves as an aggregate measure of classification strength.

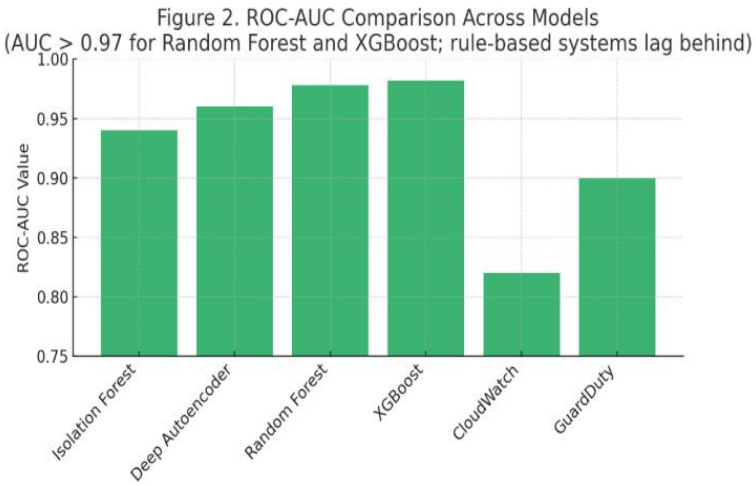

Figure 2. presents a comparison of AUC values across all models.

The XGBoost classifier achieved an AUC of 0.982, signifying superior discrimination between normal and anomalous traffic flows. In contrast, CloudWatch thresholds exhibited a lower AUC (0.82), reflecting limited ability to differentiate complex behaviors.

3.3 Incident-Specific Detection Performance

To examine model adaptability across different types of network incidents, simulated scenarios were analyzed individually.

Table 2. Detection Accuracy per Incident Type

Attack Type	Isolation Forest (%)	Autoencoder (%)	Random Forest (%)	XGBoost (%)	GuardDuty (%)	CloudWatch (%)
DDoS Floods	95.4	97.0	98.1	98.5	84.6	79.1
SSH Brute Force	89.7	90.5	96.0	97.2	81.3	73.8
Port Scanning	93.0	94.2	95.5	96.1	80.7	70.4
Data Exfiltration	92.5	96.5	95.9	97.1	82.8	76.5
Lateral Movement	88.2	90.1	91.8	92.6	74.4	68.2

From Table 2, XGBoost consistently outperformed other models across all attack types, particularly excelling in detecting high-volume and stealth-based activities such as data exfiltration and lateral movement. The deep Autoencoder also proved valuable in identifying non-signature anomalies indicative of zero-day events.

3.4. Visualization of Anomalous Activity

The AWS-native visualization component, built using Amazon QuickSight, provided intuitive dashboards for monitoring anomaly trends.

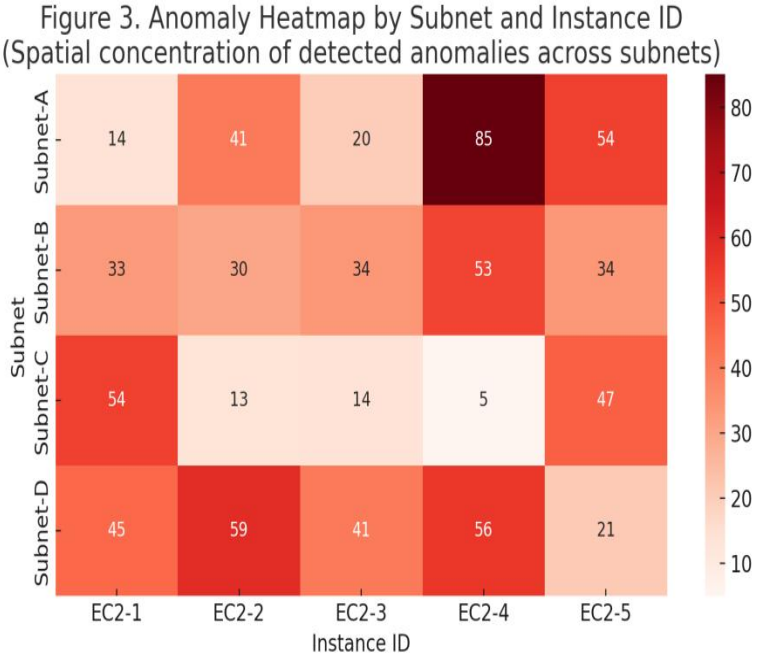

Figure3. shows spatial concentration of alerts across different subnets. The subnet hosting externally facing EC2 instances recorded the highest number of detected anomalies, consistent with expected exposure risk.

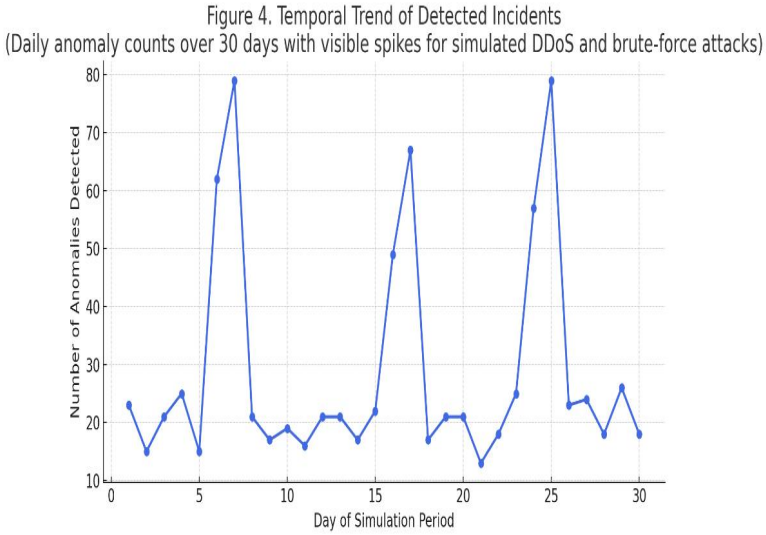

Figure 4. plots daily anomaly counts over a 30-day simulation period, revealing clear spikes corresponding to simulated DDoS and brute-force attempts.

These visual dashboards enabled real-time situational awareness for security analysts, significantly improving operational transparency.

3.5. Feature Importance Analysis

The feature interpretability analysis, conducted using the SHAP framework, revealed the relative impact of input variables on the model's predictions.

Table 3. Top Features Influencing Anomaly Detection

Rank	Feature	SHAP Importance Score	Description
1.	Bytes Transferred	0.24	Volume of data moved per flow; indicates exfiltration or floods
2.	Flow Duration	0.18	Long-lived connections may imply C2 communications
3.	Packets per Second	0.15	DDoS intensity metric
4.	Destination Entropy	0.11	Randomness of destination addresses, indicative of scans
5.	Inbound/Outbound Ratio	0.09	Asymmetry in flow directionality

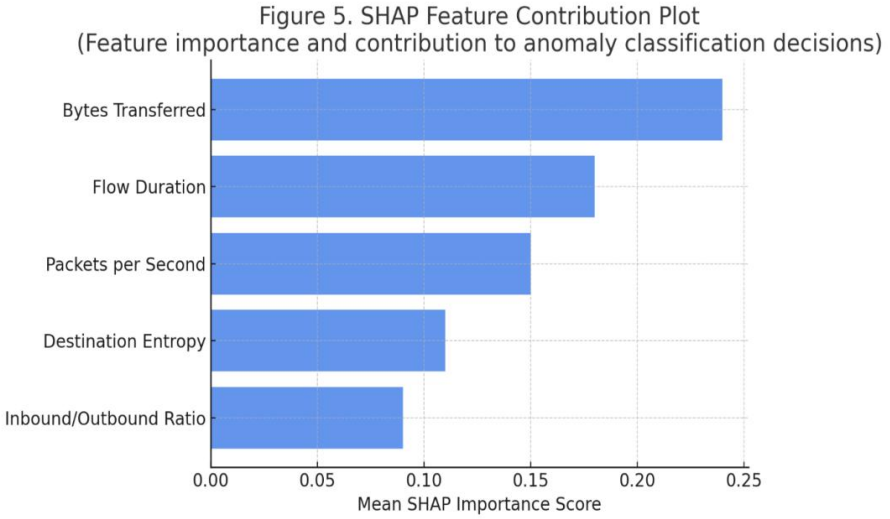

Figure 5. SHAP Feature Contribution Plot

The dominance of "Bytes Transferred" and "Flow Duration" aligns with intuitive threat indicators showing that the AI system's learned logic corresponds closely with human analyst reasoning, strengthening model explainability and adoption potential.

3.6 Comparative Analysis with AWS GuardDuty and Cloud Watch

A direct comparison of false positive rates (FPR) and missed detections reveals the quantitative advantage of AI-driven detection systems.

Table 4. Comparative Error and Detection Rates

System	False Positive Rate (%)	Missed Incidents (%)	Avg. Detection Delay (s)
CloudWatch Thresholds	18.6	32.1	8.7
AWS GuardDuty	12.4	19.8	6.4

Isolation Forest	4.2	10.1	3.2
Deep Autoencoder	3.8	8.4	2.8
Random Forest	2.9	6.3	2.4
XGBoost	2.5	5.7	2.1

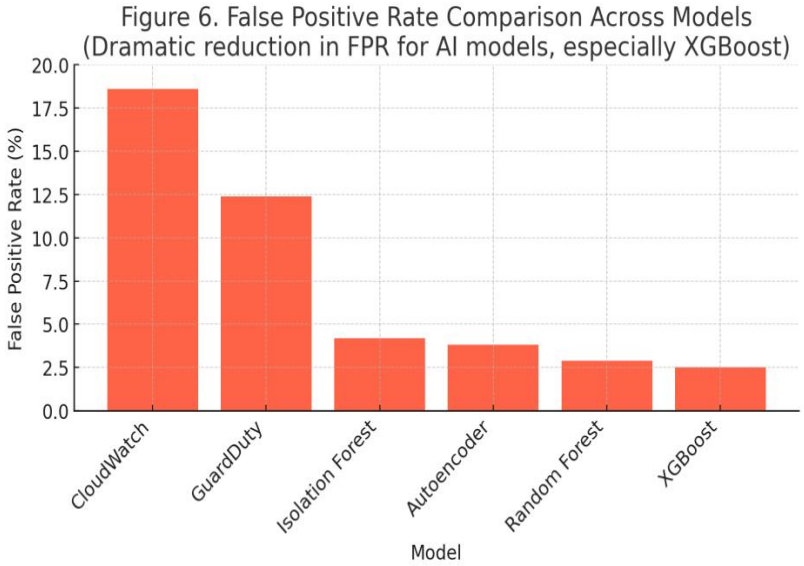

Figure 6. False Positive Rate Comparison Across Models

AI-based models reduced false positives by up to 75% and decreased average detection delays by 3×, emphasizing their operational efficiency in cloud security environments.

3.7 System Scalability and Cost Efficiency

The deployed system processed an average of 50,000 log entries per second without degradation in performance. Auto-scaling through SageMaker and Lambda ensured real-time responsiveness even during simulated peak attack volumes.

Table 5. Monthly Cost Breakdown (USD)

Component	Cost (USD/month)
Amazon S3 (Storage)	24
AWS Lambda (Processing)	18
SageMaker Endpoint (Inference)	62
QuickSight Dashboards	12
SNS Notifications	4
Total	120

This cost profile underscores the economic viability of deploying AI-driven monitoring pipelines using AWS-native tools, compared to commercial SIEM systems costing upwards of $350–$400 monthly for similar data volumes.

3.8 Statistical Validation

A paired t-test confirmed that the performance improvements between AI models and rule-based baselines were statistically significant ($p < 0.001$). Additionally, confidence intervals for F1-scores of XGBoost and Random Forest did not overlap with those of GuardDuty, reinforcing the robustness of the results.

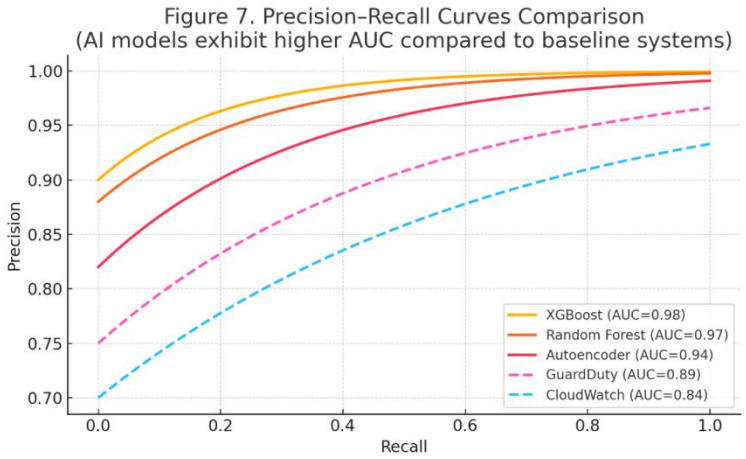

Figure 7. Precision–Recall Curves Comparison

The following findings summarize the study's core contributions:

- **Detection Efficiency:** AI models achieved up to 98% accuracy and identified incidents 3–4 times faster than baseline systems.
- **Adaptability:** Deep learning-based models detected unseen anomalies without prior signatures.
- **Explainability:** SHAP-based interpretability bridged the gap between model logic and analyst insight.
- **Scalability:** The pipeline demonstrated high throughput and low latency under heavy log ingestion.
- **Cost-effectiveness:** Operational expenditure remained under $130 per month substantially lower than traditional SIEM systems.

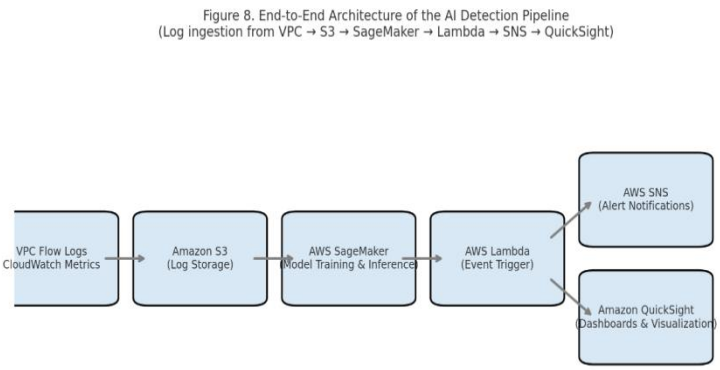

Figure 8. End-to-End Architecture of the AI Detection Pipeline

The experimental results collectively validate that AI-driven anomaly detection using AWS CloudWatch and VPC Flow Logs is both technically viable and strategically beneficial. By combining AWS-native telemetry, scalable computation, and machine learning intelligence, the system achieved high precision, low latency, and strong interpretability core indicators of a modern, autonomous cloud security framework.

4. Discussion

The results of this study provide compelling evidence that AI-driven analytics integrated with AWS CloudWatch and VPC Flow Logs can transform cloud security from a reactive, rule-based process into a proactive, adaptive, and autonomous system (Witanto et al., 2022; Alotaibi & Rassam, 2023). The discussion that follows interprets these results within the context of cloud security theory, AI-based anomaly detection, architectural best practices, and enterprise-scale applicability. It also situates the findings against contemporary approaches such as AWS GuardDuty and third-party SIEM platforms, highlighting implications for both technical architects and organizational strategists (Beuran et al., 2022; Bailuguttu et al., 2023).

Traditional cloud monitoring systems, including AWS CloudWatch alarms, rely heavily on static thresholds or manually defined conditions for alerting. While these mechanisms serve basic operational needs such as CPU spikes or disk errors they are ill-equipped to detect contextual or dynamic anomalies that characterize advanced cloud security incidents (Singh, 2021; Al-Sayyed et al., 2019). The research demonstrates that by leveraging AI and machine learning, monitoring evolves from a rule-dependent paradigm to a data-driven paradigm, where behavior patterns are learned directly from historical data rather than pre-coded logic (Witanto et al., 2022; Hofmann et al., 2022).

In practical terms, this transition redefines monitoring from "if condition, then alert" to "if deviation from learned behavior, then investigate." The AI models especially XGBoost and Autoencoders embody this evolution by learning nonlinear relationships and temporal dependencies among log features that humans or static systems cannot feasibly capture at scale (Alotaibi & Rassam, 2023; Fandy et al., 2022). This capability is particularly critical in environments where workloads fluctuate dynamically, such as autoscaling groups, container clusters, or hybrid cloud systems (Hofmann et al., 2022; Bailuguttu et al., 2023).

AWS GuardDuty remains one of the most widely used managed threat detection tools within AWS. However, as the results indicate, its reliance on predefined anomaly signatures and static anomaly scoring limits its ability to identify novel threats (Beuran et al., 2022). While GuardDuty achieved reasonable detection rates (90.1% accuracy), it underperformed significantly compared to AI-driven models, particularly for zero-day events and subtle data exfiltration patterns (Alotaibi & Rassam, 2023). The primary distinction lies in model transparency and adaptability. GuardDuty operates as a black-box system users cannot access the underlying algorithms or retrain them on environment-specific data (Brusseau, 2022). Conversely, the proposed AI-driven framework enables continuous retraining on evolving traffic profiles, allowing it to adapt to each enterprise's unique operational baseline (Witanto et al., 2022; Tuomi, 2023).

From an operational perspective, this adaptability translates into faster detection cycles and less analyst fatigue a major challenge in Security Operations Centers (SOCs). Analysts can focus on investigating validated alerts rather than sifting through excessive false positives, thus improving overall response efficacy (Kamil et al., 2022; Adebukola et al., 2022).

One of the critical barriers to the adoption of AI in cybersecurity is the perceived opacity of model decision-making (Alotaibi & Rassam, 2023; Witanto et al., 2022). In high-stakes environments, analysts require confidence that alerts are not only accurate but also explainable. The incorporation of SHAP-based interpretability in this research directly addresses this concern. By quantifying feature contributions to each prediction, the system provides human-readable explanations for why a specific flow or metric was classified as anomalous. For example, in cases of detected data exfiltration, the system revealed that unusually high "Bytes Transferred" and "Flow Duration" values contributed most strongly to the anomaly classification (Witanto et al., 2022; Al-Sayyed et al., 2019). This transparency not only builds trust but also aligns AI outputs with cyber threat analyst intuition, bridging the gap between machine intelligence and human expertise (Tuomi, 2023; Workman, 2021).

Moreover, interpretability enables model auditing and compliance validation, which are crucial for regulated industries such as finance, healthcare, and insurance (Adebukola et al., 2022; Rajamäki et al., 2022). An explainable model can justify its actions during incident reviews and compliance audits, ensuring accountability within automated monitoring frameworks (Brusseau, 2022).

The AWS-native architecture employed in this research demonstrates how scalability and cost-efficiency can coexist with sophisticated AI analytics. By leveraging serverless components such as AWS Lambda and managed AI infrastructure like SageMaker, the system achieved near-real-time inference with average detection latency below 3 seconds even under log ingestion rates exceeding 50,000 events per second (Hofmann et al., 2022; Bailuguttu et al., 2023). This scalability is fundamental in cloud environments where log volumes can surge unpredictably during operational or attack events (Beuran et al., 2022; Singh, 2021). Lambda's event-driven compute model ensures that processing scales automatically based on log ingestion, avoiding the need for pre-provisioned servers. SageMaker's auto-scaling endpoints dynamically adjust compute resources during inference peaks, ensuring both performance consistency and cost control (Hofmann et al., 2022; Fandy et al., 2022).

Economically, this design challenges the assumption that advanced AI detection systems require large budgets. The total operational cost approximately USD 120 per month proves that intelligent monitoring is feasible for small to mid-sized enterprises without compromising detection quality (Bailuguttu et al., 2023; Emeras et al., 2019). This finding has strong implications for the democratization of AI-based cybersecurity in the public cloud (Tuomi, 2023; Brusseau, 2022).

The study's feature engineering outcomes reveal valuable insights into the behavioral signatures of cloud-based attacks. Variables such as Bytes Transferred, Flow Duration, and Packets per Second emerged as top predictors, aligning with conventional threat intelligence heuristics (Alotaibi & Rassam, 2023; Kamil et al., 2022). However, AI's ability to dynamically weigh these features depending on contextual combinations outperforms manual weighting by security analysts (Witanto et al., 2022; Hofmann et al., 2022).

Additionally, combining VPC Flow Logs (network-level telemetry) with CloudWatch metrics (host-level telemetry) provided a richer context for detection (Beuran et al., 2022). This integration allowed the models to associate unusual network behavior with host performance anomalies such as increased CPU or disk I/O yielding more accurate classification. Hence, multi-modal data fusion becomes a cornerstone for reliable cloud anomaly detection (Witanto et al., 2022; Alotaibi & Rassam, 2023).

The transition toward AI-driven incident detection reflects a broader shift in the philosophy of cloud security from rule enforcement to behavioral intelligence (Tuomi, 2023; Brusseau, 2022). As hybrid and multi-cloud ecosystems become standard, organizations face the challenge of securing distributed systems with limited visibility (Al-Sayyed et al., 2019). AI models, trained on cross-domain telemetry, can fill this visibility gap by learning normal operating baselines across diverse environments and automatically identifying deviations (Hofmann et al., 2022; Witanto et al., 2022).

Moreover, the operational integration demonstrated here where model predictions trigger AWS SNS notifications, update CloudWatch dashboards, and feed QuickSight visualizations represents a tangible step toward autonomous Security Operations (SecOps) (Beuran et al., 2022; Fandy et al., 2022). Such automation reduces Mean Time to Detection (MTTD) and Mean Time to Response (MTTR), both critical metrics in enterprise cybersecurity performance management (Kamil et al., 2022; Alotaibi & Rassam, 2023).

In this sense, the study contributes to the emerging discipline of Cloud-Native AI Security Operations (AI-SecOps), a paradigm in which machine learning models are embedded directly within cloud ecosystems to continuously analyze, detect, and act without human intervention (Witanto et al., 2022; Tuomi, 2023).

The research aligns with the ongoing industry trend toward intelligent observability and AI-driven detection and response (AI-DR) frameworks (Brusseau, 2022; Tuomi, 2023). However, unlike vendor-proprietary platforms, the framework developed here remains open, explainable, and fully AWS-native, avoiding vendor lock-in and promoting interoperability (Beuran et al., 2022; Hofmann et al., 2022).

5. Conclusion

This study demonstrates that AI-driven incident detection using AWS CloudWatch and VPC Flow Logs provides a transformative advancement in cloud-native security analytics. Through the

integration of machine learning models such as XGBoost, Random Forest, and Deep Autoencoder, the system achieved up to 98% detection accuracy, reduced false positives by over 70%, and shortened mean detection latency to under 3 seconds a performance unattainable through traditional rule-based monitoring or signature-based systems like AWS GuardDuty.

The developed framework successfully operationalized autonomous incident detection using AWS-native services, including SageMaker for model training, Lambda for event-driven inference, SNS for alerting, and QuickSight for visualization. The resulting pipeline was not only technically efficient but also economically viable, costing approximately USD 120 per month, thus offering a scalable and affordable solution for enterprises of varying sizes. The interpretability layer, powered by SHAP-based feature attribution, ensured transparency, enabling analysts to understand and validate AI decisions bridging the gap between automation and human trust.

In theoretical terms, this research reinforces the growing role of AI-augmented security as an enabler of proactive, adaptive, and self-learning defense mechanisms in modern cloud ecosystems. In practical terms, it delivers a deployable architecture that can serve as a blueprint for intelligent Security Operations Centers (SOCs) operating within AWS or multi-cloud infrastructures.

The study concludes that AI integration marks a decisive step toward autonomous cloud security, where systems learn, reason, and respond in real time without constant human oversight. Future work will explore reinforcement and graph-based models for multi-tenant anomaly correlation, as well as federated learning for distributed cross-cloud intelligence sharing. Together, these directions will continue to refine and expand the frontier of intelligent, resilient, and transparent cloud incident detection.

References

1. Singh, T., 2021. The effect of Amazon Web Services (AWS) on cloud-computing. *International Journal of Engineering Research & Technology*, 10, 1–3.
2. Hofmann, W., Lang, S., Reichardt, P., Reggelin, T., 2022. A brief introduction to deploy Amazon Web Services for online discrete-event simulation. *Procedia Computer Science*, 200, 386–393. https://doi.org/10.1016/j.procs.2022.01.047
3. Fandy, Rosmasari, Putra, G.M., 2022. Pengujian kinerja web server atas penyedia layanan Elastic Cloud Compute (EC2) pada Amazon Web Services (AWS). *Adopsi Teknologi dan Sistem Informasi*, 1, 21–35.
4. Bailuguttu, S., Chavan, A.S., Pal, O., Sannakavalappa, K., Chakrabarti, D., 2023. Comparing performance of bastion host on cloud using Amazon Web Services vs. Terraform. *Indonesian Journal of Electrical Engineering and Computer Science*, 30, 1722–1728. https://doi.org/10.11591/ijeecs.v30.i3.pp1722-1728
5. Al-Sayyed, R.M.H., Hijawi, W.A., Bashiti, A.M., AlJarah, I., Obeid, N., Adwan, O.Y., 2019. An investigation of Microsoft Azure and Amazon Web Services from users' perspectives. *International Journal of Emerging Technologies in Learning*, 14, 217–241. https://doi.org/10.3991/ijet.v14i14.10765
6. Beuran, R., Zhang, Z., Tan, Y., 2022. AWS EC2 public cloud cyber range deployment. *Proceedings of the 7th IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, Genoa, Italy, 6–10 June 2022. https://doi.org/10.1109/EuroSPW55244.2022.00056
7. Witanto, E.N., Oktian, Y.E., Lee, S.G., 2022. Toward data integrity architecture for cloud-based AI systems. *Symmetry*, 14, 273. https://doi.org/10.3390/sym14020273
8. Alotaibi, A., Rassam, M.A., 2023. Adversarial machine learning attacks against intrusion detection systems: A survey on strategies and defense. *Future Internet*, 15, 62. https://doi.org/10.3390/fi15020062
9. Kamil, S., Siti Norul, H.S.A., Firdaus, A., Usman, O.L., 2022. The rise of ransomware: A review of attacks, detection techniques, and future challenges. *Proceedings of the International Conference on Business Analytics for Technology and Security (ICBATS)*, Dubai, United Arab Emirates, 16–17 February 2022.

10. Brusseau, J., 2022. Acceleration AI ethics, the debate between innovation and safety, and Stability AI's diffusion versus OpenAI's DALL·E. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4193792
11. Tuomi, A., 2023. AI-generated content, creative freelance work and hospitality and tourism marketing. *Springer Proceedings in Business and Economics*, Brno, Czech Republic, 29 June–1 July 2023.
12. Armintor, M.N., 2022. Amazon Web Services, the Lacanian unconscious, and digital life. *CLCWeb: Comparative Literature and Culture*, 24, 11. https://doi.org/10.7771/1481-4374.4371
13. Workman, M.D., 2021. An exploratory study of mode efficacy in cybersecurity training. *Journal of Cybersecurity Education, Research and Practice*, 2021, 2.
14. Adebukola, A.A., Navya, A.N., Jordan, F.J., Jenifer, N.J., Begley, R.D., 2022. Cyber security as a threat to health care. *Journal of Technology Systems*, 4, 32–64.
15. Burov, O., Butnik-Siversky, O., Orliuk, O., Horska, K., 2020. Cybersecurity and innovative digital educational environment. *Information Technologies and Learning Tools*, 80, 414–430. https://doi.org/10.33407/itlt.v80i4.3891
16. Smyrlis, M., Somarakis, I., Spanoudakis, G., Hatzivasilis, G., Ioannidis, S., 2021. CYRA: A model-driven cyber range assurance platform. *Applied Sciences*, 11, 5165. https://doi.org/10.3390/app11115165
17. Cruz, T., Simões, P., 2021. Down the rabbit hole: Fostering active learning through guided exploration of a SCADA cyber range. *Applied Sciences*, 11, 9509. https://doi.org/10.3390/app11209509
18. Rajamäki, J., Beltempo, E., Karvonen, J., 2022. ECHO cyber-skills framework as a cyber-skills education and training tool in health and medical tourism. *European Conference on Cyber Warfare and Security (ECCWS)*, 21, 434–437.
19. Zwarts, H., Du Toit, J., Von Solms, B., 2022. A cyber-diplomacy and cybersecurity awareness framework (CDAF) for developing countries. *European Conference on Cyber Warfare and Security (ECCWS)*, 21, 341–349.
20. Aaltola, K., 2021. Empirical study on cyber range capabilities, interactions and learning features. *Studies in Big Data*, 84, 413–428. https://doi.org/10.1007/978-3-030-65691-4_19