Mahlaki Henry

Mount St.  Security Operations Center

4 May 2025

# SOC Machine Learning Rec

## Introduction

Security Operations Centers (SOCs) operate in an ever-evolving threat landscape where past strategies quickly become obsolete. Existing tools are static, requiring frequent reconfiguration yet often yielding suboptimal results. This research explores the intersection of cybersecurity and machine learning to identify adaptive strategies and tools that enhance SOC infrastructure and align with its evolving goals. For the past 10 weeks, I have been researching trends in Artificial Intelligence (AI) and Machine Learning (ML) within SOCs. Additionally, I have surveyed current SOC analysts at Mount St. Mary's University (MSMU) to understand their operations. This outline proposes assessing whether ML is necessary for the MSMU SOC and provides an overview of the initial standards for integrating AI.

## Limitations

Security Operations Centers' challenges today extend beyond merely adapting to technological trends and enhancing existing tools. These limitations are rooted in a strategic necessity, highlighting the imperative for SOCs to evolve in the age of rapid AI advancements. AI should not be viewed solely as a tool for modernization; instead, it must be recognized as a critical response mechanism to emerging AI-driven threats that could compromise security integrity.  Embracing AI in this capacity is essential for SOCs to stay ahead of sophisticated threats and to ensure they can maintain complete defense mechanisms in an increasingly complex cyber landscape.

## Proposed Approach

To combat these challenges, we aim to enhance our capabilities by integrating ML models with our cybersecurity tools. Modern security operations face significant challenges due to the vast amounts of data and quickly evolving attack techniques, which traditional rule-based systems struggle to manage effectively. Many of the systems in our toolkit can incorporate ML, enabling us to implement advanced techniques such as firewall behavior analysis, email phishing detection, and improved intelligence gathering for Security Information and Event Management (SIEM).

A case study conducted by Komandina et al. explored the use of an Anomaly Detection model with firewall data. Their research demonstrated that the model uncovered subtle deviations in network behavior that traditional detection methods might overlook. Similarly, our current approach builds on these principles, showcasing that machine learning-enhanced systems outperform traditional methods by facilitating adaptive learning and enabling quicker response times.

To enhance the discussion surrounding our case study, it makes sense to begin with a smaller dataset before tackling the extensive firewall data. This practical approach not only allows us to train machine learning models but also offers a way to validate their outputs in a manageable context. By doing so, we can gradually integrate these models into our workflow, which will help us understand the semantics of our network traffic more effectively. It’s important to acknowledge that integrating external tools within a structured environment can present significant challenges. Therefore, ensuring a smooth and seamless integration process is essential for success.

In parallel with these considerations, we can look to advancements in the field, such as Google Chronicle. This cloud-native security operations platform combines Security Information and Event Management (SIEM), Security Orchestration, Automation, and Response (SOAR), and advanced threat intelligence into a cohesive solution. Built on Google Cloud's infrastructure, Chronicle employs AI-driven analytics, including Gemini's natural language processing, to automate critical aspects such as threat detection, investigation, and response. By unifying SIEM and SOAR workflows in a single interface, it eliminates the need to switch between different tools, consolidating alerts into contextualized cases that are enriched with AI-generated summaries and remediation recommendations.

With these insights, we can propose a strategic plan aimed at gradually enhancing our Security Operations Center (SOC) capabilities. Instead of overhauling our existing infrastructure, our focus should be on improving and building upon it. A key element of this enhancement will be to strengthen the connection between our SIEM system and SOAR tool. This initiative will ultimately lead to improved threat detection and response effectiveness through the incorporation of machine learning.

# Model Selection

## Overview

To address the idea of a constantly changing environment, it is essential to implement a system that adapts dynamically. This is where Machine Learning (ML) comes in. ML-driven automation can prioritize and contextualize both new and familiar alerts, improving threat detection and response capabilities.

I propose the following models to achieve this task: OpenSearch is an open-source search and analytics suite derived from Elasticsearch. It supports various pre-trained ML models for semantic search and clustering tasks. Similarly, H20.ai's H20 AutoML is an open-source platform that automates the machine learning process, including training and tuning models tailored to specific datasets. SpaCy, an open-source natural language processing library, provides pre-trained models and pipelines for various linguistic tasks. Utilizing these readily available models allows seamless integration within existing infrastructure and tools.

## Log Analysis

OpenSearch features a powerful pre-trained model, "huggingface/sentence-transformers/all-MiniLM-L6-v2," designed for effective log analysis. This model employs semantic search to cluster various threat patterns based on keywords found in logs. As a sentence transformer, it converts raw, unlabeled logs into vectors, allowing for k-NN search to identify semantically similar events. For example, if a faculty member who typically sends 10-15 emails per day suddenly triggers over 500 emails in an hour, which may appear as normal SMTP traffic in traditional log analysis, this unusual behavior can be flagged as anomalous by the model. By transforming raw email server logs into vectors, the model enables k-NN searches to identify this behavior as semantically akin to known cases of compromised accounts or internal threats. Implementing this model enhances our Security Information and Event Management (SIEM) system, improving threat detection and revealing hidden complexities.

## Contextual Enrichment

Before the Ingestion phase into the Security Orchestration, Automation, and Response (SOAR) from the SIEM, a step called contextual enrichment is performed to enhance alerts with additional context. This process takes the log information that triggered the alert, as predefined by the SIEM, and parses it to retrieve meaningful metadata (e.g., Alert Type, Email Domain, IP, Port). SpaCy's Named Entity Recognition (NER) model converts unstructured logs into structured fields. These structured fields will serve as feature inputs for the model, alongside the alert, enabling the model to predict the necessary actions.

## Alert Classification

Subsequently, H2O.ai's H2O AutoML will be trained on the context-enriched alerts generated through the SIEM to accurately classify responses to those alerts. H2O AutoML automates the process of training and tuning multiple machine learning models (e.g., XGBoost, Generalized Linear Models, Deep Learning) within a user-defined time frame or model count limit. It generates a leaderboard ranking models by performance metrics (e.g., AUC, Log Loss, RMSE). This approach allows users to clean data, select the best-performing model from the leaderboard, and deploy it while also considering continuous learning through analyst feedback.

## Feedback Loop

Fortunately, H2O.ai also provides a technique for explaining AI decisions. H2O's Model Explainability Interface, designed to be user-friendly and automatic, produces global explanations (i.e., how changes in one feature impact overall model predictions) and local explanations (i.e., how specific feature values influence individual predictions) for each H2O model. Most explanations are visual, facilitating quick pattern recognition to describe the model's actions both globally and locally.

## Evaluation Metric

It’s essential to evaluate the effectiveness of any ML-inspired system. These evaluations establish a framework that ensures everyone contributes to an accurate and adaptive system. The OpenSearch model uses metrics that balance true positives (correctly flagged threats) and false positives (benign logs misclassified as threats). High precision ensures that analysts are not overwhelmed by irrelevant alerts, while high recall reduces the risk of missing actual threats. Additionally, the Silhouette Score metric helps validate the quality of log clusters formed through k-NN. Scores closer to 1 indicate that logs with similar threat patterns are tightly grouped, whereas dissimilar logs remain distinct.

Using SpaCy's NER model, the NLP model is evaluated on precision, recall, and F1-score to ensure accurate identification of security-relevant entities. Low recall (false negatives) risks overlooking critical threats, while low precision (false positives) wastes analysts' time. The ML model developed with H2O AutoML is assessed using AUC-ROC and Precision-Recall curves to manage imbalanced data and minimize critical false negatives. Additionally, H2O.ai provides numerous web interfaces that provide centralized views of a model's performance, like H2O Model Analyzer which allows users to simulate scenarios, analyze, and evaluate a model's behavior and limitations in response to real-world changes..

Lastly, the feedback loop may be the most crucial component. Metrics like Accuracy Drift, indicating a gradual decline in the predictive performance of a machine learning model over time, and feedback coverage, which reflects the percentage of ML predictions reviewed by analysts, determine the effectiveness of the analyst validation. Without these metrics, the model would not be as adaptive or reliable.

These metrics allow us to gradually trust the model rather than granting it unchecked freedom. They help prevent "garbage in, garbage out" scenarios, ensuring that the model's precision remains reliable amid continuous adaptation.

# Outline Approach

## General Overview

The deployment begins with the SIEM, where raw security logs are ingested and preprocessed to clean, normalize, and deduplicate the data. Next, OpenSearch's ML capabilities will perform semantic searches and k-nearest neighbors (k-NN) clustering to enrich the alert data with confidence scores and additional threat context, identifying subtle patterns that traditional rule-based methods might miss.

The enriched data flows into the NLP context enrichment stage. SpaCy's custom-trained NER model extracts critical security entities, such as IP addresses and security types, from unstructured text to produce structured metadata.

This enriched metadata is then fed into the ML-enhanced SOAR, which leverages H20 AutoML to automatically train and select an ensemble model that predicts alert severity and confidence score and determines the required actions based on historical incident patterns and security type. The SOAR automation aspect receives the predictions and automatically triggers appropriate playbooks (e.g., remedial acts) based on conditional logic—while lower-generated alerts are deemed for an analyst for further review.

Finally, the process concludes with analysts reviewing and validating the automated response using H2O's explainability interface. Their feedback refines the model further and guides computerized actions (e.g., push notifications, case creation) based on the alert's rank and confidence score. This ML-enhanced approach ensures SOC effectiveness and adaptability in the face of evolving security challenges.

### Expected Outcomes

Beyond improved effectiveness and adaptability, the ML-enhanced pipeline yields several valuable outcomes. For instance, SOC can expect greater accuracy in threat detection, leading to fewer false positives and more analyst-centered risk assessments. The pipeline can also improve operations by automating routine tasks and reducing labor time in sifting through log data. Moreover, allowing a machine to learn and react based on threat patterns enhances SOC situational awareness, improving organizations' decision-making and long-term security.

## Conclusion

To wrap up, following the trend of technology converging to AI and a foundational understanding of SOC's operations. Combining an ML-enhanced SIEM with a dynamic ML-based SOAR tool will improve SOC's operations and encourage innovative thinking. The proposal addresses the need to adhere to cybersecurity's complex and rapidly evolving landscape by enhancing threat detection with ML, enabling context-aware automated responses with NLP, and maintaining model transparency through explainable AI—all while emphasizing the importance of analysts. By gradually transitioning from static rule-based systems to a dynamic, continuous learning framework, we can achieve greater outcomes than before.

## Works Cited

D, Peer. “Building a Feedback Loop System for Continuous Improvement in Machine.” Peerdh.Com, 23 Sept. 2024, peerdh.com/blogs/programming-insights/building-a-feedback-loop-system-for-continuous-improvement-in-machine-learning-models.

Deloitte. “Enhancing Soc Efficiency through Artificial Intelligence (AI …” Deloitte, June 2024, www2.deloitte.com/content/dam/Deloitte/us/Documents/risk/us-laao-whitepaper-final.pdf.

Fa, Abe. “Lime vs. Shap.” Medium, 23 May 2022, https://medium.com/@afanta/lime-vs-shap-a92623e95c4.

Fathima, Shaistha. “LIME vs SHAP: A Comparative Analysis of Interpretability Tools.” Lime vs Shap: A Comparative Analysis of Interpretability Tools, 26 Feb. 2024, www.markovml.com/blog/lime-vs-shap#choosing-the-right-tool.

Gill, Dr. Jagreet Kaur. “How Ai-Driven SOAR Platforms Enhance Incident Response.” Real Time Data and AI Company, Xenonstack Inc, 4 Mar. 2025, www.xenonstack.com/blog/enhancing-incident-response-with-ai-driven-soar-platforms.

Google. “Google Chronicle Siem and Soar for Enhanced Cybersecurity.” CyberProof, 7 Nov. 2024, www.cyberproof.com/siem/google-chronicle-siem-and-soar-for-enhanced-cybersecurity/.

Greunke, Brian. “Building AI and Machine Learning into Modern SOC Security.” Home, May 2023, www.meetascent.com/resources/ai-machine-learning-modern-soc.

H2O. “H2O AUTOML: Automatic Machine Learning.” H2O AutoML: Automatic Machine Learning - H2O 3.46.0.7 Documentation, 2024, docs.h2o.ai/h2o/latest-stable/h2o-docs/automl.html.

H2O. “Model Explainability.” *Model Explainability - H2O 3.46.0.7 Documentation*, 2024, docs.h2o.ai/h2o/latest-stable/h2o-docs/explain.html.

Hatfield, Maria. “OpenSearch as a SIEM Solution.” OpenSearch, Mar. 2025, opensearch.org/blog/OpenSearch-as-a-SIEM-Solution/.

Komadina, Adrian, et al. “Comparative Analysis of Anomaly Detection Approaches in Firewall Logs: Integrating Light-Weight Synthesis of Security Logs and Artificially Generated Attack Detection.” Sensors (Basel, Switzerland), vol. 24, no. 8, 20 Apr. 2024, p. 2636, doi:10.3390/s24082636.

McGraw, Terry. “Automating a Security Operations Center Using Machine Intelligence.” Secureworks, 6 July 2023, www.secureworks.com/blog/elevating-security-operations-with-artificial-intelligence-and-automation.

Onge, Michael St. “How We Combine AI, ML, LLMs & Human Expertise to Prioritize CNAPP Alerts.” Tamnoon, 19 Aug. 2024, tamnoon.io/blog/ai-cnapp-alert-prioritization/.

OpenSearch. “Pretrained Models.” OpenSearch Documentation, 4 Apr. 2025, opensearch.org/docs/latest/ml-commons-plugin/pretrained-models/.

Palo Alto Networks. “Artificial Intelligence and Machine Learning in the Security Operations Center.” Palo Alto Networks, May 2020, www.paloaltonetworks.com/resources/techbriefs/artificial-intelligence-and-machine-learning-in-the-security-operations-center.

Ribeiro, Marco Tulio, et al. "'why Should I Trust You?': Explaining the Predictions of Any Classifier." arXiv.org, 16 Feb. 2016, https://arxiv.org/pdf/1602.04938v1.

Sandialabs. "Sandialabs/Alert-Triage." GitHub, 14 Jan. 2016 github.com/sandialabs/alert-triage/.

SchesmuTwo. "AI-Powered SOC: Automating Incident Response with Machine Learning and SOAR Tools." Medium, 23 Sept. 2024, medium.com/@akramtalibi1902/ai-powered-soc-automating-incident-response-with-machine-learning-and-soar-tools-70ab343e9402. Accessed 3 Feb. 2025.

Smith, John &. "The Rise of AI-Powered SOC Tools: Revolutionizing Security Operations." AI in the SOC: A Guide for Security Professionals, 12 Feb. 2024, www.linkedin.com/pulse/rise-ai-powered-soc-tools-revolutionizing-security-operations-j3smc/.

Google. "Google Chronicle Siem and Soar for Enhanced Cybersecurity." CyberProof, 7 Nov. 2024, www.cyberproof.com/siem/google-chronicle-siem-and-soar-for-enhanced-cybersecurity/.

Tiwari, Sanchit. "Feedback Loop in Machine Learning – Labeling Data." LinkedIn, 4 Feb. 2021, www.linkedin.com/pulse/feedback-loop-machine-learning-labeling-data-sanchit-tiwari.

Vinod, Viji, and V Sarala Devi. "A USER-CENTRIC MACHINE LEARNING FRAMEWORK for CYBER SECURITY OPERATIONS CENTER." @International Research Journal of Modernization in Engineering, vol. 4945, 2024, www.irjmets.com/uploadedfiles/paper/issue_4_april_2024/53254/final/fin_irjmets1713462441.pdf, doi:10.56726/IRJMETS53254. Accessed 17 Feb. 2025.