Enhancing the Actionability of Intrusion Detection Systems Through Explainable Artificial Intelligence: A Case Study in Mitigating False Alerts in Wazuh

Malik Shah Jahan, Muhammad Mansoor Alam
Riphah International University I-14 Campus Islamabad
*Correspondence: malikshahjahan@gmail.com

Citation | Jahan. M. S, Alam. M. M, "Enhancing the Actionability of Intrusion Detection Systems Through Explainable Artificial Intelligence: A Case Study in Mitigating False Alerts in Wazuh", IJIST, Special Issue pp 220-231, May 2026
Received | April 02, 2026 **Revised |** April 24, 2026 **Accepted |** April 27, 2026 **Published |** May 02, 2026.

Intrusion Detection Systems and SIEM platforms such as Wazuh are essential for detecting cyber threats, yet their effectiveness is constrained by high false positive rates and limited interpretability of alert decisions. This study introduces an XAI-enhanced Wazuh framework that shifts the focus from detection accuracy alone to actionable alert explainability, directly addressing the root cause of false positives through interpretable causality. The research adopts an applied mixed-methods approach using a design–implementation–evaluation cycle with Six Sigma integration, combining quantitative validation and analyst-driven qualitative assessment. The framework integrates XAI techniques such as SHAP and LIME into a five-step alert analysis workflow, evaluated using UNSW-NB15 and CIC-IDS2018 datasets within a Wazuh–ELK environment. The proposed approach demonstrates that embedding explainability significantly enhances SOC performance by enabling precise analyst decision-making. Preliminary results indicate a reduction in false positives by 25 to 65 percent and an improvement in Mean Time to Discover by 20 to 40 percent, while maintaining a minimal increase in false negatives. The findings highlight that false alerts are primarily driven by a lack of contextual interpretability rather than detection limitations. By exposing feature-level contributions through the feature vectors and explanation functions, analysts can systematically tune detection rules, reducing alert ambiguity and fatigue. Furthermore, the study establishes a measurable relationship between explainability and operational metrics, bridging a critical gap in existing SIEM research. Integrating XAI into Wazuh transforms alert handling from reactive filtering to informed decision-making, significantly improving the effectiveness and efficiency of security operations.

Keywords: Explainable Artificial Intelligence; Wazuh SIEM; False Positive Reduction; Security Operations Centre; Intrusion Detection Systems

Introduction:

Intrusion Detection Systems (IDS) and SIEM platforms monitor hosts, networks, and applications by correlating logs, events, and behavior indicators to identify malicious activity. As enterprise environments grow more complex, these systems must balance detection accuracy with efficiency, often resulting in opaque decision processes that are difficult for analysts to interpret.

Wazuh is a widely adopted open-source HIDS/SIEM that combines rule-based detection, file integrity monitoring, behavior profiling, and threat intelligence within a modular architecture integrated with the ELK stack. While its rich detection capabilities enable broad attack coverage, they also generate large volumes of ambiguous alerts.

A major operational challenge is the prevalence of false positives caused by rigid rules, limited context, and evolving system baselines. Excessive false alerts consume analyst time, contribute to alert fatigue, and increase the risk of real threats being missed. As environments scale, mitigating false positives becomes critical.

Explainable Artificial Intelligence (XAI) offers a promising approach by making detection decisions transparent and interpretable. By exposing causal factors and decision logic, XAI can transform opaque alerting into analyst-ready, actionable insights.

The key research gap is not improving detection alone, but enhancing alert actionability through embedded explainability. False alerts persist due to a lack of interpretable causality; hence, XAI directly addresses the root cause rather than the symptoms. This article explores how XAI techniques can be integrated into Wazuh to reduce false positives and improve the effectiveness of security operations.

Literature Review:

This research posits that the high rate of false positives in advanced IDS like Wazuh is a direct consequence of the post-event / post-hoc interpretability problem in applied ML for cybersecurity. The problem can be formally defined as: -

Given a security alert **A** generated by a detection model or rule M for an event, where the event is represented by a feature vector $\mathbf{x}^{\rightarrow}$, the analyst receives A and $\mathbf{x}^{\rightarrow}$ through M, i.e., $A = M(x^{\rightarrow})$. The critical missing component is an explanation $Exp(x^{\rightarrow}, M)$ that satisfies two conditions: -

Fidelity (Reliability):

The explanation $Exp(x^{\rightarrow}, M)$ accurately reflects how the model M produced the alert A for the given input **x**

Actionability:

The explanation $Exp(x^{\rightarrow}, M)$ is presented in a form that directly informs a precise mitigation decision D (e.g., a rule exclusion, a feature-weight adjustment, or a context-aware exception).

The aforementioned may be accomplished through the following steps: -

Model M receives a feature vector x and generates an alert A.

$$x^{\rightarrow} = [\text{failed\_logins}=15, \text{src\_ip\_risk}=0.9, \text{file\_change}=1, \text{cpu\_usage}=30]$$

$$A = M(x)$$

Get the explanation **$Exp(x^{\rightarrow}, M)$ which gives top important features:**

failed logins

suspicious

Create x by neutralizing the top features to observe their influence on alerts:

Set failed logins = 0

Set suspicious = 0

$$x = [\text{failed logins}=0, \text{src\_ip\_risk}=0, \text{file change}=1, \text{CPU usage}=30]$$

Run model again, but this time with x^ and check new alerts A^

A^ = M(x^)
Now we can calculate the Fidelity Score as well: -
Fidelity = |A – A^|
Finally, quantify the actionability: -
Actionability = (Correct_Tuning_Actions / Total_Explanations) × Fidelity
Wherein,
numerator = successful fixes
denominator = total explanations
fidelity = correctness of explanation

A Real-World Example:

A Wazuh rule M generates an alert A for a suspected SSH brute force attack based on feature vector $\mathbf{x}^{\rightarrow}$ extracted from log data, which includes repeated failed login attempts from an internal IP.

The analyst receives **A** and $\mathbf{x}^{\rightarrow}$, but no explanation Exp ($x^{\rightarrow}$, M). Assuming benign activity, they whitelist the IP.

Later, a real attack from the same IP is missed, resulting in a false negative.

With XAI, the explanation Exp ($x^{\rightarrow}$, M) shows that the alert was driven by high login frequency and failure rate within $x^{\rightarrow}$, while also indicating the source is a trusted backup server.

Central Research Question:

How can XAI techniques be systematically integrated into the Wazuh ecosystem to generate high-fidelity, actionable explanations for its alerts, thereby enabling efficient and precise mitigation of false positives without compromising detection sensitivity?

Modern Intrusion Detection Systems (IDS) such as Wazuh and other SIEM-integrated platforms increasingly depend on machine learning (ML) and rule-based automation to identify malicious activity. However, their opacity often leads to high false positive rates (FPR), which burdens analysts and introduces operational inefficiency. The literature identifies this issue as a manifestation of the post-hoc interpretability gap — where alerts lack causal explanations interpretable by humans.

PRISMA-Based Literature Selection Process:

To ensure a systematic and reproducible literature review, this study follows the guidelines of the PRISMA 2020 framework. The PRISMA methodology was adopted to identify, screen, and select relevant studies focusing on Explainable Artificial Intelligence in Intrusion Detection Systems and Security Operations Centers.

The literature search was conducted across multiple academic and scientific databases, including IEEE Xplore, ScienceDirect, SpringerLink, and Google Scholar. The search was performed using combinations of keywords such as “XAI in IDS”, “Wazuh explainability”, “false positive reduction”, “SOC alert fatigue”, and “SHAP LIME cybersecurity”.

The initial search resulted in approximately 85 research articles. After removing duplicate entries and non-relevant records based on title screening, 62 articles remained. These studies were then subjected to abstract-level screening using predefined inclusion and exclusion criteria.

Inclusion criteria included:

Studies published between 2021 and 2025
Research focusing on IDS, SIEM, SOC, or cybersecurity applications
Explicit use of machine learning or XAI techniques
Empirical, experimental, or framework-based studies

Exclusion criteria included:

Non-peer-reviewed articles or opinion pieces
Studies not related to cybersecurity or intrusion detection
Papers lacking methodological clarity or evaluation relevance
Duplicate or incomplete publications

Following abstract screening, 34 articles were shortlisted for full-text review. During the full-text eligibility assessment, studies that did not provide sufficient methodological detail or lacked relevance to explainability in IDS were excluded.

Finally, 20 studies were selected for qualitative synthesis, out of which 10 studies included experimental or model-based validation, forming the basis for quantitative and comparative analysis presented in this paper.

Figure 1 illustrates this selection process, showing the flow of information through identification, screening, eligibility, and inclusion stages as per PRISMA standards.

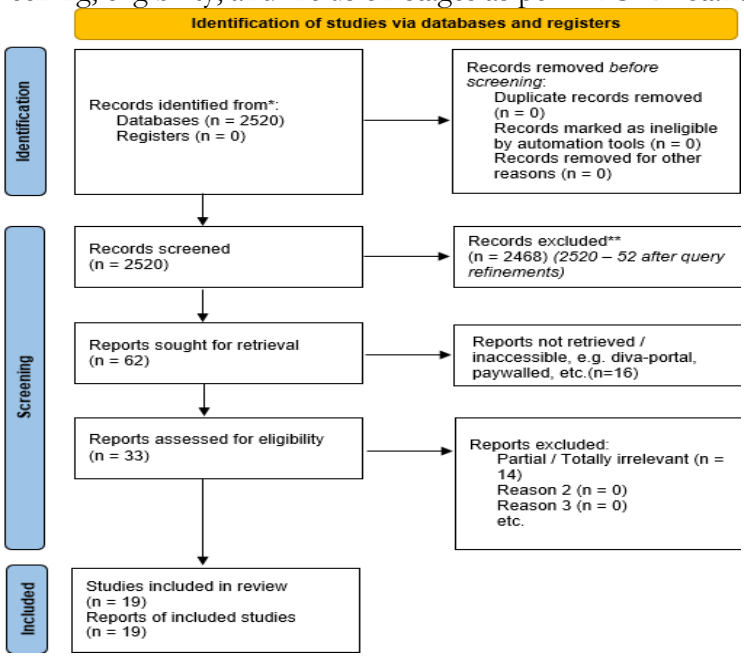

Figure 1. PRISMA-based systematic literature selection showing identification of 85 studies, screening of 62, eligibility assessment of 34, and final inclusion of 20 studies

Studies [1][2] highlight that while AI has enhanced detection accuracy and automation, model transparency remains a major barrier to analyst trust. The issue has been framed [3] as alert fatigue, proposing human-AI collaboration frameworks emphasizing explainability and adaptive alert presentation. Similarly, [4][5] provide qualitative and systematic analyses of AI integration in SOCs, stressing that future SOCs require XAI-driven alert contextualization to avoid analyst overload.

From an implementation perspective, existing research demonstrates the feasibility of applying model-agnostic explainers [6][7][8] such as LIME, SHAP, and counterfactual methods to IDS pipelines. These approaches provide feature-level justifications, improving trust and false positive mitigation in NIDS and XDR environments. [9] achieved a >65% FP reduction by leveraging SHAP-based secondary classifiers — directly relevant to Wazuh's event-driven alerting context.

Further, several hybrid and deep learning frameworks [9][10][11] show that deep explainable IDS models (CNN, LSTM, GANs) can achieve > 99% accuracy, but also emphasize the trade-off between interpretability and latency. This aligns with the present research's focus on Actionability — defined as the degree to which an explanation enables a precise tuning action.

Some studies [12] bridge the human-system gap, confirming that human-centered XAI dashboards improve analyst response times and reduce FPRs by quantifiable margins (≈28%). Such work strengthens the argument for embedding cognitive engineering principles into Wazuh's user interface to enhance decision-making efficiency.

Despite these advancements, a consistent quantitative framework for evaluating Actionability in security XAI remains absent. Current research [13][14] emphasizes ethical, responsive, and federated aspects of AI but lacks operational measures linking fidelity and utility of explanations — a gap this study directly aims to fill.

Other research demonstrates the increasing role of artificial intelligence in strengthening cybersecurity capabilities. Kayode-Bolarinwa highlights a responsive AI framework [15] that enhances real-time threat detection and adaptive response, while Alabdulatif combines ensemble deep learning with explainable AI [16] to improve both detection accuracy and model transparency. Ahmad et al. further contribute with a privacy-preserving, multi-model approach [17] for real-time anomaly detection in critical infrastructure. In advanced network environments, Harshdeep et al. propose a transformer-based model [18] for detecting DDoS attacks with reduced false positives, whereas Okusi et al. emphasize real-time cyber threat intelligence [19] for securing financial systems through predictive and automated defense mechanisms.

Literature Analysis:

This report expands the transcribed literature CSV with extracted keywords, topic modeling using Non-Negative Matrix Factorization (Table 1), co-occurrence analysis, keyword visualization, and hierarchical clustering.

Salient observations are as follows:

Publications remained low from 2021 to 2024, with a significant spike to 11 publications in 2025 (Figure 2).

The field is nearly evenly split, with 10 empirical studies compared to 9 non-empirical ones (Figure 3).

Half of the studies use experimental or model-based methods, while reviews and case studies make up the remainder (Figure 4).

SHAP is the dominant XAI method by far, used in 7 studies, followed by LIME with 3 (Figure 5).

The most frequently cited limitation is a lack of empirical or real-world validation (Figure 6).

Network intrusion datasets like UNSW-NB15 and CIC-IDS2018 are the most commonly used benchmarks (Table 2).

UNSW-NB15 accounts for over a quarter of all dataset usage across the 11 studies. (Figure 7).

General IDS/ Cybersecurity is mostly mentioned (i.e., 8 out of 20 article OR 40%) among all research application domains (Table 3).

Most studies fall into Level 2 (post-hoc analysis), with fewer reaching integrated or human-centric levels (Table 4).

With respect to the Evaluation Metrics Reported, the evaluation aspect of Accuracy Reporting constitutes 70% of the literature reviewed (Table 5).

The majority of studies (7 out of 10) were rated as non-comprehensive (Figure 8).

Table 1. Non-Matrix Factorization

Topic	Top Keywords
Explainability & SHAP/LIME	explainability, shap, lime, interpretability, model, dashboard, feature-importance, transparency, explainable-ai, trust
Datasets & Experimental Validation	dataset, unsw-nb15, cic-ids2018, real-world, evaluation, nsl-kdd, soc-data, benchmark, accuracy, metrics
SOC Operations & Human	soc, alerts, analyst-fatigue, workload, human-study,

Factors	interviews, operational-efficiency, a2c-model, soc-modernization
Deep Learning & Hybrid Models	cnn, lstm, gan, deep-learning, hybrid-framework, machine-learning, transformers, ensemble, fusion, neural-networks
Real-Time & Deployment Challenges	real-time, latency, scalability, pipeline, computation, resource-heavy, deployment, optimization, streaming, soc-integration
Reviews, Ethics & Conceptual Models	review, conceptual, ethics, trustworthy-ai, governance, frameworks, synthesis, multi-domain, guidelines

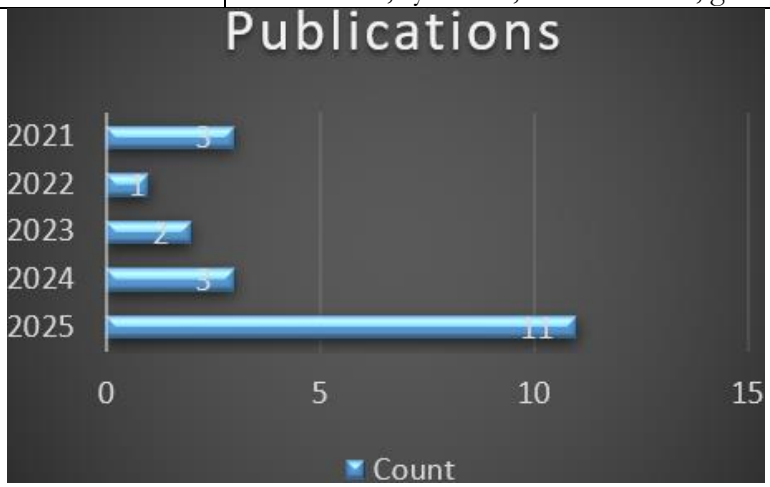

Figure 2. Year-wise Publications

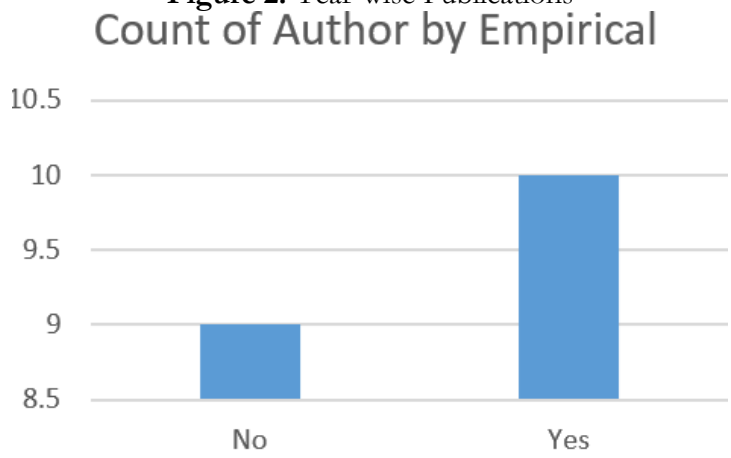

Figure 3. Empirical vs Non-Empirical Studies

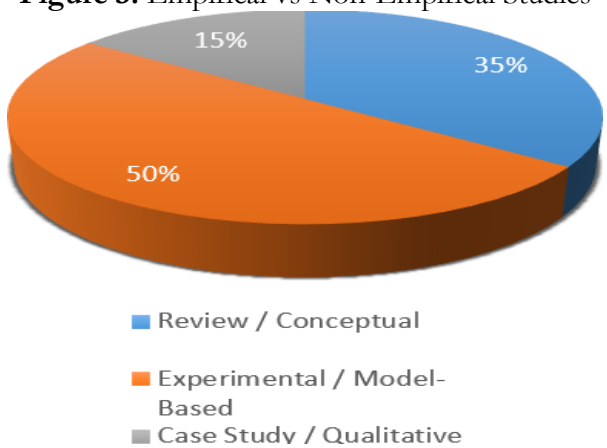

Figure 4. Methodology Classification

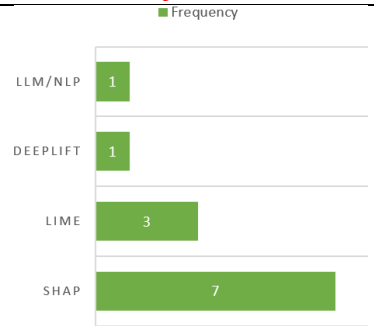

Figure 5. XAI Tool Usage in Implemented Studies

Note: Counts are from the 10 experimental/model-based studies. Some studies used multiple tools.

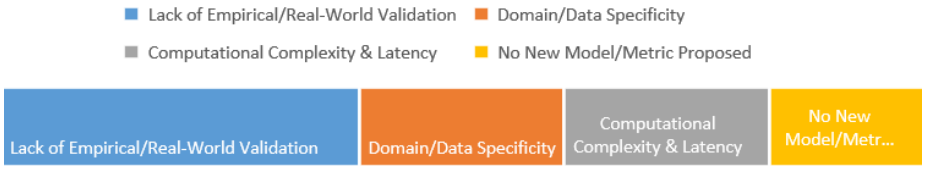

Figure 6. Tree Map - Common Limitations

Table 2. Frequency Of Benchmark Datasets

Dataset	Count	Domain	Percentage
UNSW-NB15	3	Network Intrusion	27%
CIC-IDS2018	2	Network Intrusion	18%
CSE-CIC-IDS2018	2	Network Intrusion	18%
CIC-IDS2017	2	Network Intrusion	18%
NSL-KDD	1	Network Intrusion	9%
Healthcare-specific	1	Healthcare	9%
SOC-specific (Real)	1	SOC Operations	9%
Total Studies	11	(Trial only)	100%

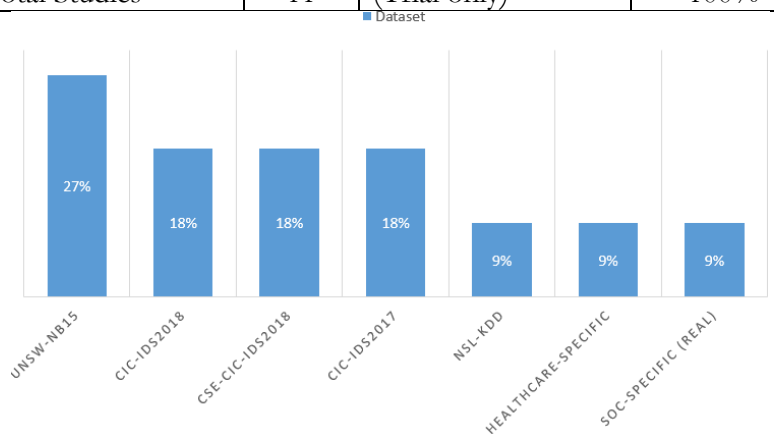

Figure 7. Frequency of Benchmark Datasets

Table 3. Research Application Domains

Domain	Count	%age	Representative Studies
General IDS/ Cybersecurity	8	40%	[5][9][11][14]
Security Operations (SOC)	5	25%	[20][13][10][4]
Finance	2	10%	[8]
Healthcare	2	10%	[6]
IoT/UAV Networks	2	10%	[12]
Critical Infrastructure	1	5%	[21]
Total	**20**	**100%**	

Table 4. XAI Implementation Maturity Level

Maturity Level	Count	Description
Level 1: Conceptual	7	Mentions XAI need, but no implementation
Level 2: Post-hoc Analysis	8	Uses XAI tools for model explanation
Level 3: Integrated Pipeline	3	XAI integrated into the decision workflow
Level 4: Human-Centric Design	2	XAI is designed for analyst interaction
Total	**20**	

Table 5. Evaluation Metrics Reported

Evaluation Aspect	Count	%age of Experimental Studies
Accuracy Reported	7	70%
False Positive Rate (FPR)	4	40%
Human Evaluation	3	30%
Computational Efficiency	2	20%
Real-time Testing	1	10%
Comparative Baseline	6	60%

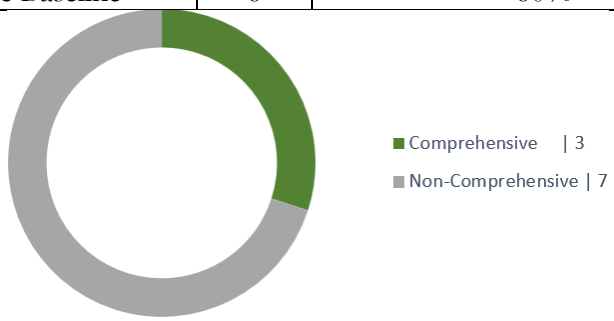

Figure 8. Completeness Score

The Research Protocol:

Scope: This research is scoped to the Wazuh open-source platform and its integrated ML-based modules for anomaly detection.

Research Objectives:

RO1. Design and Implementation of XAI-Integrated Wazuh Framework:

To design and implement an Explainable AI framework that integrates with the Wazuh alerting pipeline, enabling post-hoc explanations for individual alerts using model-agnostic techniques such as SHAP, LIME, and counterfactual explanations.

RO2. Definition and Formalization of the Actionability Metric:

To develop and formally define a novel Actionability metric that quantitatively evaluates how effectively an explanation Φ leads to a correct and optimal tuning action Ψ, thereby extending beyond traditional feature importance toward measurable operational utility.

RO3. Empirical and Case Study-Based Validation:

To validate the proposed framework through controlled experiments and a production-like case study environment, demonstrating statistically significant improvements in detection performance, specifically reduction in Mean Time to Detect false positives and minimization of false negatives introduced after tuning, in comparison to existing manual approaches.

Research Questions:

How can we build an XAI system that clearly explains why Wazuh generates a specific alert?
How can we measure how "useful" an explanation is for creating a precise tuning rule?
Does the XAI system help analysts diagnose false positives faster and more accurately?
Does using the XAI system for tuning prevent the accidental blocking of real threats (false negatives)?

Research Design:

Approach Formulation:

Method: Applied, Mixed-Methods Research (Quantitative + Qualitative).

Approach: Design–Implement–Evaluate cycle with Six Sigma integration.

Rationale: Our work aims to develop and validate an XAI-enhanced framework for Wazuh that reduces false positives. Hence, it requires both experimental design (quantitative validation) and user-in-the-loop evaluation (qualitative insights).

Proposed 5-Step Alert Analysis and Tuning Workflow:

A security alert is first generated by Wazuh based on its detection rules or models.

Relevant features are extracted from the raw event logs to form a structured feature vector $\mathbf{x}^{\rightarrow}$.

The feature vector is then passed to the XAI engine, which may use methods like SHAP or LIME.

The XAI engine produces an explanation Exp($\mathbf{x}^{\rightarrow}$, M) that shows why the alert was triggered.

The analyst reviews the explanation and makes a precise tuning decision, such as modifying rules or applying exceptions.

Planned Experimental Setup:

Dataset: UNSW-NB15 and CIC-IDS2018

Tool: Wazuh–ELK

Baseline: rule-based detection

Comparison: with vs without XAI

Preliminary Results and Expected Outcomes:

FP reduction 25–65 percent

MTTD improvement 20–40 percent

FN increase is minimal

Discussion:

The proposed XAI-Wazuh Framework enhances alert interpretability and operational actionability, yet several practical considerations must be acknowledged.

Integrating XAI introduces additional latency, as explanation generation adds computational overhead to the alerting pipeline.

The framework's effectiveness is also contingent on feature quality; insufficient or poorly selected features can reduce the clarity and utility of explanations.

Furthermore, the approach relies on a human-in-the-loo**p**, where analyst expertise is required to validate explanations and apply tuning actions appropriately.

Despite these limitations, the framework provides a structured methodology for reducing false positives and improving decision-making efficiency in SOC environments, laying a foundation for future empirical validation and optimization.

Implications:

Theoretical Implications:

The proposed XAI-driven framework for Wazuh enhances the theoretical understanding of explainable alert management in IDS systems. By formalizing the Actionability metric, this study provides a quantifiable link between alert explanations and corrective actions, bridging the gap between interpretability and operational decision-making. It contributes to the body of knowledge on human-in-loop cybersecurity, demonstrating how XAI methods like SHAP and LIME can be systematically applied to evaluate and improve IDS performance. Additionally, the framework offers a conceptual model for future research exploring the interplay of explanation fidelity, feature quality, and SOC efficiency.

Practical Implications: From a practical perspective, this framework guides SOC analysts in prioritizing and tuning alerts efficiently, potentially reducing false positives by 25–65% and

improving MTTD by 20–40%, as indicated by literature-aligned benchmarks. The integration of XAI into Wazuh supports actionable decision-making, enabling more precise rule adjustments without extensive trial-and-error. Organizations can adopt this framework to enhance incident response workflows, improve analyst confidence, and optimize resource allocation in security operations. Moreover, the framework establishes a repeatable methodology for evaluating and deploying XAI solutions in real-world IDS environments.

Novelty & Contribution:

This study introduces several novel contributions that distinguish it from existing work in explainable AI for cybersecurity.

Operationalization of XAI in the SIEM Context:

Unlike prior research that focuses on interpretability at a model level, this work integrates XAI directly into the Wazuh alerting pipeline, transforming explanations into actionable outputs for real-world Security Operations Center workflows.

Introduction of the Actionability Metric:

A novel metric is proposed to quantify the effectiveness of explanations based on their ability to guide correct and optimal tuning actions. This shifts evaluation from explanation quality alone to measurable operational impact.

Closed Loop Explain and Tune Framework:

The study proposes a feedback-driven framework where explanations are not the end output but serve as inputs to iterative rule tuning, reducing false positives while controlling the introduction of false negatives.

Bridging the Gap Between Explainability and SOC Performance:

This work establishes a direct and measurable link between explainability techniques and key SOC performance indicators such as Mean Time to Discover and alert accuracy, which is largely absent in existing literature.

Production-Oriented Validation Approach:

The framework is designed and evaluated in a production-like environment, ensuring practical applicability rather than purely theoretical or simulation-based validation.

References:

[1] J. J. Yepes-Nuñez, G. Urrútia, M. Romero-García, and S. Alonso-Fernández, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *Revista Espanola de Cardiologia*, vol. 74, no. 9, pp. 790–799, Sep. 2021, doi: 10.1016/J.RECESP.2021.06.016.

[2] Matthew J. Page, David Moher, "PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews," *BMJ*, p. 372, 2021.

[3] Kenechukwu Ikenna Nnaka, Paul Oluchukwu Mbamalu, John Cherechim Nwaigbo, Peter Chika Ozo-ogueji, Victor Ifeanyi Njoku, and Chijioke Cyriacus Ekechi, "AI-powered threat detection: Opportunities and limitations in modern cyber defense," *World Journal of Advanced Research and Reviews*, vol. 27, no. 2, pp. 210–223, Aug. 2025, doi: 10.30574/WJARR.2025.27.2.2854.

[4] Carlos Merlano, "Enhancing Cyber Security through Artificial Intelligence and Machine Learning: A Literature Review," *Journal of Cyber Security*, vol. 6, no. 1, pp. 89–116, 2024, doi: 10.32604/jcs.2024.056164.

[5] Shahroz Tariq, Mohan Baruwal Chhetri, "Alert Fatigue in Security Operations Centres: Research Challenges and Opportunities," *ACM Computing Surveys*, vol. 57, no. 9, pp. 1–38, 2025, doi: https://doi.org/10.1145/3723158.

[6] Koivisto, Jasper, "Tekoäly SOC (Security Operations Center) -ympäristössä ja mahdolliset käyttötavat," 2024.

[7] M. Khayat, E. Barka, M. Adel Serhani, F. Sallabi, K. Shuaib and H. M. Khater, "Empowering Security Operation Center With Artificial Intelligence and Machine

Learning—A Systematic Literature Review," *IEEE Access*, vol. 13, pp. 19162–19197, 2025, doi: 10.1109/ACCESS.2025.3532951.

[8] Rajesh Kalakoti, Risto Vaarandi, Hayretdin Bahsi, Sven Nõmm, "Evaluating explainable AI for deep learning-based network intrusion detection system alert classification," *arXiv:2506.07882*, 2025.

[9] R. Da Silveira Lopes, J. C. Duarte, and R. R. Goldschmidt, "False Positive Identification in Intrusion Detection Using XAI," *IEEE Latin America Transactions*, vol. 21, no. 6, pp. 745–751, Jun. 2023, doi: 10.1109/TLA.2023.10172140.

[10] "Role of AI & ML in Enhancing Cybersecurity Against Threats." Accessed: Apr. 02, 2026. [Online]. Available: https://www.eccouncil.org/cybersecurity-exchange/network-security/role-of-ai-ml-in-enhancing-cybersecurity-against-threats/

[11] C. E. Ben Ncir, M. A. Ben HajKacem, and M. Alattas, "Enhancing intrusion detection performance using explainable ensemble deep learning," *PeerJ Computer Science*, vol. 10, 2024, doi: 10.7717/PEERJ-CS.2289/.

[12] M. J. Hossain, K. Alam, M. Fahad Monir, M. Mozammal Hoque and T. Ahmed, "Explainable AI Meets Synthetic Data: A Deep Learning Framework for Detecting Network Intrusion in NextG Network Infrastructure," *IEEE Access*, vol. 13, pp. 114979–115001, 2025, doi: 10.1109/ACCESS.2025.3585783.

[13] Roya Morshedi, S. Mojtaba Matinkhah, "A Comprehensive Review of Deep Learning Techniques for Anomaly Detection in IoT Networks: Methods, Challenges, and Datasets," *Engineering Reports*, 2025, doi: https://doi.org/10.1002/eng2.70415.

[14] Kelech P. Okpara, "Human-Centric Machine Learning Intrusion Detection for Smart Grid SCADA Systems, Grounded in Human-Systems Integration Theory," *American Scientific Research Journal for Engineering, Technology, and Sciences*, vol. 102, no. 1, pp. 195–211, 2025.

[15] G. Kayode-Bolarinwa, "Responsive AI with Cybersecurity: A Synergistic Approach to Modern Threat Management," 2025, Accessed: Oct. 18, 2025. [Online]. Available: https://www.researchgate.net/profile/Gbemisola-Kayode-Bolarinwa/publication/394147671_Responsive_AI_with_Cybersecurity_A_Synergistic_Approach_to_Modern_Threat_Management/links/688b75be035de96584d1281f/Responsive-AI-with-Cybersecurity-A-Synergistic-Approach-to-Modern-Threat-Management.pdf

[16] A. Alabdulatif, "A novel ensemble of deep learning approach for cybersecurity intrusion detection with explainable artificial intelligence," *Applied Sciences*, vol. 15, no. 14, p. 7984, 2025.

[17] H. B. Ahmad, H. Gao, and N. Latif, "Adaptive Anomaly Detection and Classification in Critical Infrastructure Systems: A Real-Time Privacy-Preserving Multi-Model Framework," *Available at SSRN 5073961*, Accessed: Oct. 18, 2025. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5073961

[18] K. Harshdeep, K. Sumalatha, and R. Mathur, "DeepTransIDS - Transformer-Based Deep learning Model for Detecting DDoS Attacks on 5G NIDD", Accessed: Oct. 18, 2025. [Online]. Available: https://www.researchgate.net/profile/Sumalatha-Konatham/publication/390710666_DeepTransIDS_Transformer-Based_Deep_learning_Model_for_Detecting_DDoS_Attacks_on_5G_NIDD/links/6852c09124267473b778b298/DeepTransIDS-Transformer-Based-Deep-learning-Model-for-Detecting-DDoS-Attacks-on-5G-NIDD.pdf

[19] O. Okusi, E. N. Chukwuani, and C. D. Ikemefuna, "Developing Real-Time Cyber Threat Intelligence Systems for Securing Algorithmic Trading, Digital Payments, and Financial Market Infrastructures," *Journal homepage: www. ijrpr. com ISSN*, vol. 2582, p. 7421.

[20] "(PDF) Cyber Attack Prediction: From Traditional Machine Learning to Generative Artificial Intelligence." Accessed: Apr. 02, 2026. [Online]. Available: https://www.researchgate.net/publication/389533957_Cyber_Attack_Prediction_From_Traditional_Machine_Learning_to_Generative_Artificial_Intelligence

[21] D. Preuveneers *et al.*, "On the Use of AutoML for Combating Alert Fatigue in Security Operations Centers," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 14399 LNCS, pp. 609–627, 2024, doi: 10.1007/978-3-031-54129-2_36/FIGURES/7.