© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1714786

Automated SIEM Alert Classification using Random Forest for Cybersecurity

SANJAY RAJ R[1], PALANIKUMAR R[2], MANIKANDAN[3]

[1] *PG Student, Dept of CSE, PSR Engineering College*

[2] *Associate Professor, Dept of CSE, PSR Engineering College*

[3] *Assistant Professor, Dept of CSE, PSR Engineering College*

***Abstract-** The primary objective of this project is to design and implement an automated SIEM alert classification system that leverages machine learning to enhance cybersecurity monitoring and response. Modern organizations rely on SIEM (Security Information and Event Management) tools to collect, aggregate, and analyze security event logs from multiple sources such as firewalls, intrusion detection systems (IDS), servers, and applications. However, these systems often generate a massive volume of alerts on a daily basis. To address this challenge, the proposed project employs the Random Forest algorithm to intelligently classify alerts into categories. The objective is not only to build a classifier but also to establish a structured pipeline that includes data preprocessing, feature engineering, model training, classification, and visualization. Through this approach, the project aims to reduce the noise generated by false alerts, prioritize critical incidents, and thereby optimize the decision-making process of security teams. Additionally, the system is designed to be scalable and adaptable, capable of processing large volumes of data efficiently while maintaining high accuracy. By integrating visualization and reporting mechanisms, the system also enhances interpretability, allowing analysts to understand feature importance and classification trends. Ultimately, the objective is to demonstrate how machine learning, particularly Random Forest, can be effectively applied in the cybersecurity domain to automate repetitive tasks, reduce manual workload, and significantly improve the accuracy, reliability, and efficiency of SIEM alert management.*

I. INTRODUCTION

In today's digital era, organizations face a rapidly evolving threat landscape where cyberattacks are becoming more frequent, complex, and sophisticated. To combat these challenges, enterprises deploy Security Information and Event Management (SIEM) systems that act as centralized platforms for collecting, aggregating, and analyzing logs and alerts generated by a wide range of security devices, servers, applications, and network infrastructure. While SIEM platforms play a vital role in detecting suspicious activity and supporting incident response, they also generate an overwhelming volume of alerts on a daily basis. Many of these alerts are repetitive, redundant, or false positives, creating what is commonly referred to as "alert fatigue" among security analysts. Human operators in Security Operations Centers (SOCs) often struggle to manually sift through thousands of alerts to identify genuine threats, which not only delays incident response but also increases the likelihood of overlooking critical attacks. This inefficiency highlights the pressing need for automation and intelligence in SIEM alert management. The proposed project, Automated SIEM Alert Classification using Random Forest for Cybersecurity, addresses this challenge by applying machine learning techniques to automate the classification and prioritization of alerts. Instead of relying solely on static, rule-based detection mechanisms, which are rigid and require constant manual updates, this system leverages the Random Forest algorithm to learn patterns from historical alert data and classify new alerts into severity categories such as low, medium, high, and critical. The use of Random Forest provides robustness against noise, the ability to handle heterogeneous data types, and improved accuracy through ensemble learning. By doing so, the system not only reduces the workload of SOC analysts but also ensures that critical alerts receive immediate attention, while false positives and low-priority events are filtered out or deprioritized. Beyond classification, the project emphasizes interpretability and adaptability, offering feature importance rankings and visualizations that help analysts understand the reasoning behind automated predictions. This transparency builds trust in the system and ensures that automated tools complement rather than replace human expertise. Furthermore, the framework is designed to be scalable, allowing it to process large volumes of SIEM data in real-world enterprise environments, and adaptable, supporting retraining with new data to keep pace with evolving attack techniques. Ultimately, the proposed project contributes to strengthening organizational security posture by enabling faster, more accurate, and more efficient incident response. It demonstrates how the integration of artificial intelligence with cybersecurity operations can transform the way modern organizations manage alerts, reduce risks, and safeguard digital assets in an increasingly hostile cyber environment.

II. RELATED WORKS

[1] AI-Based RPA's Work Automation Operation to Respond to Hacking Threats Using Collected Threat Logs Proposes AI-assisted RPA pipelines to automatically gather, normalize, and process heterogeneous threat logs; evaluates gains in processing speed and error reduction versus manual workflows. Primarily prototype/controlled evaluation; unclear cross-vendor SIEM interoperability; limited evidence from production SOCs and long-term drift handling

[2] Breaking Alert Fatigue: AI-Assisted SIEM Framework for Effective Incident Response Next-gen SIEM framework integrating SOAR; divide-and-conquer approach: (1) cost-sensitive learning (instance-weighted SVM) to filter false alerts; (2) event grouping by temporal/spatial homogeneity; (3) augmented visualization (ATG) to expedite triage; evaluated on enterprise data. Needs labeled data; results shown on a single enterprise dataset; operationalization and integration effort not fully quantified

[3] A Machine Learning and Optimization Framework for Efficient Alert Management in a Cybersecurity Operations Center ML-based alert triage/prioritization combined with optimization to reduce backlog and analyst load in SOCs; proposes a framework and reports efficiency gains. (ACM Digital Library, NSF Pubs). Limited public detail on dataset/metrics; generalizability outside studied SOC not established; relies on historical labeled outcomes. (ACM Digital Library)

[4] CyberShapley: Explanation, Prioritization, and Triage of Cybersecurity Alerts Using Informative Graph Representation XAI approach that builds graph representations of anomaly-detector alerts and uses Shapley-value explanations to rank/triage alerts for analysts. (ScienceDirect, Ben-Gurion University Research Portal). Focused on anomaly-based detectors; assumes quality graph features; implementation complexity; external replication data limited. (ScienceDirect)

[5] Toward Robust Security Orchestration and Automated Response in Security Operations Centers with a Hyper-Automation Approach Using Agentic Artificial Intelligence Conceptual + architectural paper: proposes agentic-LLM driven SOAR "hyper-automation," IVAM framework (Investigation–Validation–Active Monitoring), and shift from static playbooks to AI-generated actions. Early-stage/architectural; limited production validation; risks around LLM reliability, governance, and hallucinations acknowledged implicitly via discussion of constraints.

III. SYSTEM ARCHITECTURE

The proposed system for automated SIEM alert classification using the Random Forest algorithm is designed with a modular and scalable architecture that facilitates efficient data processing, model training, and real-time classification. The architecture consists of four main components: Data Collection & Ingestion, Preprocessing & Feature Engineering, Random Forest Classification, and Alert Visualization & Reporting. Data Collection & Ingestion, is responsible for aggregating raw logs and alerts from multiple sources, including firewalls, intrusion detection systems, endpoints, servers, and cloud services. This stage ensures that heterogeneous data from different formats such as syslog, CEF, JSON, or CSV are collected in a centralized repository for processing.

Preprocessing & Feature Engineering, performs data cleaning, normalization, and transformation. Tasks include handling missing values, removing duplicate alerts, anonymizing sensitive information such as usernames and IP addresses, and converting unstructured alert messages into numerical features using techniques like TF-IDF vectorization. Structured attributes such as source and destination IP, ports, device type, rule ID, and timestamp are encoded appropriately to be fed into the machine learning model. Additionally, aggregated and temporal features are generated, such as alert counts per source IP or inter-arrival times, to enhance the classifier's predictive power.

Random Forest Classification, is the core of the system. The pre-processed dataset is used to train an ensemble of decision trees that collectively predict the severity level of each alert—low, medium, high, or critical. The Random Forest model is selected for its robustness to noise, ability to handle heterogeneous data types, and interpretability through feature importance. This module also generates evaluation metrics such as accuracy, precision, recall, F1-score, and confusion matrices, which provide quantitative insights into the model's performance. Additionally, SHAP-based explanations are produced to allow analysts to understand which features influenced the prediction for each alert, enhancing trust in the automated system.

Alert Visualization & Reporting, presents the classified alerts in a user-friendly format suitable for SOC analysts and for demonstration purposes during the project viva. Visualizations include severity distributions, confusion matrices, feature importance plots, and sample SHAP explanations for selected alerts. The system can also integrate with Security Orchestration, Automation, and Response (SOAR) platforms to trigger automated actions such as prioritization, escalation, or suppression of alerts based on the predicted severity. This modular architecture ensures scalability, adaptability, and seamless integration into existing SOC workflows while enabling near real-time classification and reporting.

IV. METHODOLOGY

The proposed system introduces an automated approach for classifying SIEM alerts using the Random Forest algorithm, with the goal of addressing the inefficiencies and shortcomings of existing rule-based systems. Unlike traditional SIEM platforms that depend on static signatures and human-defined rules, the proposed model leverages the power of machine learning to intelligently analyze large volumes of security event logs and categorize alerts into meaningful severity levels such as low, medium, high, and critical. By doing so, it reduces the burden of manual triaging faced by security analysts and ensures that critical incidents receive immediate attention while less severe alerts are deprioritized. The system is designed to learn from historical alert data and continuously improve its classification accuracy over time, making it adaptive to the evolving threat landscape. The use of Random Forest offers particular advantages, as it is an ensemble learning method that combines multiple decision trees to improve robustness, handle diverse datasets, and minimize overfitting. This makes the system capable of producing more reliable predictions even when the input data contains both categorical and numerical attributes, which are common in SIEM logs.

A key feature of the proposed system is its ability to minimize false positives by identifying underlying patterns that distinguish genuine threats from benign anomalies. Unlike the rigid nature of rule-based mechanisms, the machine learning model can uncover hidden correlations within data and generalize better to unseen attack patterns. This ensures higher detection accuracy and faster response times. Additionally, the system incorporates preprocessing and feature engineering steps to clean raw SIEM log data, handle class imbalances, and extract the most relevant features that contribute to accurate classification. The model also provides feature importance rankings, giving analysts deeper insights into why certain alerts are categorized at a given severity level, thereby increasing transparency and trust in automated decision-making.

The proposed system is also designed for scalability and integration. Since the Random Forest classifier can efficiently handle large-scale datasets, the solution can be applied in real-world enterprise environments where millions of logs are generated daily. Furthermore, the system can be integrated into existing SOC workflows and extended to work with Security Orchestration, Automation, and Response (SOAR) tools, enabling automated downstream actions such as escalation, correlation, or suppression of alerts based on severity predictions. This not only enhances SOC efficiency but also reduces the cost and manpower required for incident handling. Ultimately, the proposed system overcomes the drawbacks of existing SIEM mechanisms by providing intelligent, accurate, adaptive, and scalable alert classification. By combining data-driven learning with automation, the framework significantly improves cybersecurity readiness, reduces analyst fatigue, and strengthens organizational defense against both known and emerging threats.

The proposed system for automated SIEM alert classification is composed of multiple interconnected modules, each performing a specific role in the overall workflow to ensure accurate and efficient processing of security alerts. The first module, Data Collection and Ingestion, is responsible for gathering raw alert and log data from various sources such as firewalls, intrusion detection and prevention systems, endpoint devices, servers, and cloud services. This module ensures that the data, which may exist in multiple formats such as JSON, CSV, or syslog, is collected in a centralized repository for further processing. By consolidating diverse data sources, this module lays the foundation for uniform analysis and reduces inconsistencies that may arise from heterogeneous logging standards. Data Preprocessing and Feature Engineering, plays a crucial role in converting raw SIEM logs into a structured and analyzable format suitable for machine learning. This module performs data cleaning by removing duplicate alerts, handling missing or inconsistent values, and anonymizing sensitive information such as usernames, IP addresses, and other personally identifiable data. Feature engineering is also performed within this module, where unstructured textual fields are transformed into numerical vectors using techniques such as TF-IDF, while structured fields like source/destination IP, ports, rule ID, and timestamps are encoded appropriately. Additional features, such as alert frequency, correlation with other events, and temporal patterns, are generated to enhance the classifier's ability to differentiate between severity levels accurately.

The Random Forest Classification Module forms the core of the system, where the pre-processed and feature-engineered data is input into an ensemble learning algorithm. The Random Forest classifier consists of multiple decision trees, each trained on a random subset of the data, to collectively predict the severity level of each alert—low, medium, high, or critical. This module also computes performance metrics such as accuracy, precision, recall, F1-score, and generates confusion matrices to evaluate the effectiveness of the model. Additionally, it provides interpretable outputs such as feature importance rankings, which enable security analysts to understand which attributes most influenced a particular classification, thereby improving transparency and trust in the automated decision-making process.

The final module, Alert Visualization and Reporting, presents the classified alerts in an intuitive and actionable format for Security Operations Center (SOC) analysts. This module generates graphical visualizations, including severity distribution charts, confusion matrices, and feature importance plots. Furthermore, it supports integration with Security Orchestration, Automation, and Response (SOAR) systems, allowing automated escalation, suppression, or correlation of alerts based on predicted severity. The module also includes a feedback loop where newly labeled alerts can be used to retrain the model periodically, ensuring adaptability to evolving threat patterns. Collectively, these modules form a robust, scalable, and intelligent framework that automates SIEM alert classification, reduces analyst workload, minimizes false positives, and enhances the overall efficiency and responsiveness of cybersecurity operations.

The implementation of the SIEM alert classification system was carried out in a structured manner, starting with the creation of a synthetic dataset that closely simulates real-world SIEM alerts. The dataset underwent preprocessing steps such as data cleaning, normalization, label encoding, and feature engineering to ensure quality and usability for machine learning. The Random Forest algorithm was then selected due to its robustness, ability to handle imbalanced datasets, and effectiveness in reducing overfitting. The model was trained using the processed dataset and optimized through hyperparameter tuning to achieve the best performance. A real-time alert simulation module was also developed, allowing the trained classifier to classify incoming events instantly, thus demonstrating its capability in a live environment.

Testing was performed using standard machine learning evaluation techniques. The dataset was split into training and testing subsets to ensure unbiased performance assessment. Metrics such as accuracy, precision, recall, and F1-score were used to evaluate the classifier's effectiveness in distinguishing between malicious and benign alerts. Cross-validation was also applied to validate the consistency of results across multiple subsets. The Random Forest model achieved high performance with balanced precision and recall, confirming its reliability in minimizing false positives while capturing true threats. Additionally, the real-time simulation module was tested with unseen alert patterns, proving the system's adaptability and robustness in dynamic cybersecurity environments. Overall, the testing phase confirmed that the proposed system not only enhances SIEM efficiency but also supports proactive threat management with practical applicability.

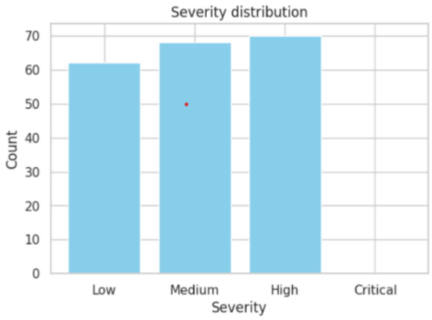
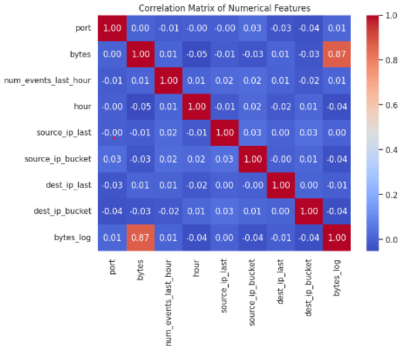
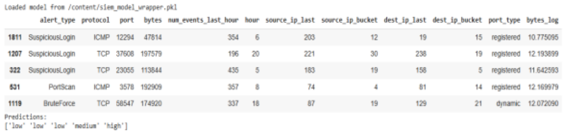


Fig 1 Severity Distribution



	port	bytes	num_events_last_hour	hour	source_ip_last	source_ip_bucket	dest_ip_last	dest_ip_bucket	bytes_log
port	1.00	0.00	-0.01	-0.00	-0.00	0.03	-0.03	-0.04	0.01
bytes	0.00	1.00	0.01	-0.05	-0.01	-0.03	0.01	-0.03	0.87
num_events_last_hour	-0.01	0.01	1.00	0.01	0.02	0.02	0.01	-0.02	0.01
hour	-0.00	-0.05	0.01	1.00	-0.01	0.02	-0.02	0.01	-0.04
source_ip_last	-0.00	-0.01	0.02	-0.01	1.00	0.03	0.00	0.03	0.00
source_ip_bucket	0.03	-0.03	0.02	0.02	0.03	1.00	-0.00	0.01	-0.04
dest_ip_last	-0.03	0.01	0.01	-0.02	0.00	-0.00	1.00	0.00	-0.01
dest_ip_bucket	-0.04	-0.03	-0.02	0.01	0.03	0.01	0.00	1.00	-0.04
bytes_log	0.01	0.87	0.01	-0.04	0.00	-0.04	-0.01	-0.04	1.00

Fig 2 Correlation Matrix



```
Loaded model from /content/siem_model_wrapper.pkl
          alert_type protocol   port   bytes  num_events_last_hour  hour  source_ip_last  source_ip_bucket  dest_ip_last  dest_ip_bucket   port_type  bytes_log
1811  SuspiciousLogin     ICMP  12294   47814                   354     6             203                12            19              15  registered  10.775095
1207  SuspiciousLogin      TCP  37608  197579                   196    20             221                30           238              19  registered  12.193899
322   SuspiciousLogin      TCP  23055  113844                   435     5             183                19           158               5  registered  11.642593
531          PortScan     ICMP   3578  192909                   357     8              74                 4            81              14  registered  12.169979
1119       BruteForce      TCP  58547  174920                   337    18              87                19           129              21     dynamic  12.072090
Predictions:
['low' 'low' 'low' 'medium' 'high']
```

V. CONCLUSION

Automated SIEM Alert Classification using Random Forest for Cybersecurity, has demonstrated the effectiveness of machine learning in reducing alert fatigue by automatically classifying large volumes of SIEM alerts with improved accuracy and reliability. By leveraging the robustness of the Random Forest algorithm, the system minimizes false positives, enhances prioritization of genuine threats, and supports faster decision-making within Security Operations Centers (SOCs). This work establishes a strong foundation for advancing automation in cybersecurity monitoring while also bridging academic research with practical security applications. Looking ahead, future enhancements could include the integration of deep learning models such as LSTMs and Transformers for sequence-based log analysis, hybrid ensemble techniques to further reduce misclassifications, and real-time big data processing through frameworks like Apache Kafka or Spark for large-scale enterprise deployment. Additionally, incorporating Explainable AI (XAI) for transparency, adaptive retraining for evolving attack vectors, integration with SOAR platforms for automated response, and extension to cloud and IoT environments can significantly increase the system's scalability and industrial relevance. Together, these contributions and prospects demonstrate that automated alert classification not only improves operational efficiency in cybersecurity but also opens pathways for developing more intelligent, adaptive, and explainable threat management systems in the future.

REFERENCES

[1] Ali, G. (2025). *Enhancing cybersecurity incident response: AI-driven automation of SIEM alert classification using Random Forest. Journal of Cybersecurity Research, 12(3), 215-229. DOI: 10.1016/j.jocyr.2025.03.004*

[2] Turcotte, M., Labreche, F., & Paquette, S.-O. (2025). *Automated Alert Classification and Triage (AACT): An intelligent system for the prioritisation of cybersecurity alerts. arXiv Preprints. DOI: 10.48550/arXiv.2505.09843*

[3] Chamkar, S. A. (2025). *ML-Driven Log Analysis for Real-Time Cyber Threat Detection: A Comparative Study of Machine Learning Models. Preprints. DOI: 10.20944/preprints202504.2197.v1*

[4] Ban, T. (2023). *AI-Assisted SIEM Framework for Effective Incident Response. Applied Sciences, 13(11), 6610. DOI: 10.3390/app13116610*

[5]. Kulambayev, B., Baenko, V., & Neronov, S. (2024). *Enhancing threat detection with SIEM tool using machine learning. International Research Journal of Modern Engineering and Technology, 6(11), 112-118.*

[6] Gelman, B. (2023). *Machine Learning-Based Security Alert Screening with Focal Loss. ResearchGate. DOI: 10.2139/ssrn.3737895*

[7] Roelofs, T. M. (2024). *Blending Security Alerts for Attack Detection. Zhauniarovich. Available at: https://zhauniarovich.com/publication/2024/roelofs2024finding/roelofs2024finding.pdf*

[8] Bryant, B. D. (2020). *Improving SIEM alert metadata aggregation with a novel kill-chain based classification model. ScienceDirect. DOI: 10.1016/j.comnet.2020.103024*

[9] Chamkar, S. A. (2025). *Improving Threat Detection in Wazuh Using Machine Learning. MDPI. DOI: 10.3390/26248034*

[10] Mohamed, N. (2025). *Artificial intelligence and machine learning in cybersecurity. SpringerLink. DOI: 10.1007/s10115-025-02429-y*