Cyber Incident Response Systems Using Machine Learning and Severity Ranking Algorithm

[1]Alabi Orobosade., [2]Ugwunna Charles., [1]Falana Olorunjube., [1]Adejimi Alaba., [1]Aborisade Dada., [1]Olakunle Abdul Sodiq

[1]Department of Computer Science, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria

[2]Department of Computer Science, Wigwe University Isiokpo, Rivers State

DOI: https://doi.org/10.51584/IJRIAS.2025.100800183

Received: 24 August 2025; Accepted: 31 August 2025; Published: 06 October 2025

ABSTRACT

Cyber incident response is an essential framework that organizations employ to effectively manage the aftermath of cyberattacks or security breaches, aiming to mitigate impacts and ensure swift recovery. Adopting Artificial Intelligence (AI) significantly advances threat detection, classification, and response efficiency. The study underscores transitioning from traditional methods to a more advanced, automated system leveraging machine learning algorithms. Various models, including Logistic Regression, Random Forest, Support Vector Classifier, Decision Tree Classifier, and Histogram Gradient Boosting Classifier, were trained and evaluated, demonstrating high accuracy in classifying network traffic and identifying cyber threats. Central to this study is the Severity Ranking Algorithm, which quantifies incident severity by integrating intensity, frequency, and potential impact, derived from a linear regression model. This algorithm enables dynamic prioritization of incidents, ensuring efficient resource allocation and timely responses. The stratification of incidents into low, medium, and high-severity categories, based on calculated severity scores, further streamlines incident response processes. The implementation highlights the effectiveness of machine learning models in enhancing cybersecurity measures. The developed Cyber Incident Response System demonstrates significant advancements in threat detection and response.

Keywords: Cyber threats, Threat Detection, Incident Response, Artificial Intelligence, Security Breaches

INTRODUCTION

Cyber incident response is a critical framework that organizations implement to manage the aftermath of a cyberattack or security breach effectively. This structured approach aims to mitigate the breach's impact, minimizing both financial losses and the time required for recovery. The responsibility to respond to cybersecurity incidents lies with the incident response (IR) function. (Naseer et al., 2021). One of the major activities in dealing with cybersecurity incidents is the collection, storage, and analysis of all data related to the incident that has happened or is happening (Naseer et al., 2021; Pierazzi et al., 2016). Although cyber incident response and recovery are crucial in most cybersecurity strategies, they are less explored than in other areas (Stave et al., 2022). Adopting Artificial Intelligence (AI) to refine incident response processes marks a significant evolution, transforming the traditional landscape into a more advanced and efficient system. Moreover, AI-driven systems continuously leverage past incidents' data to refine their detection algorithms, ensuring an ever-improving defense mechanism. (Cranford, 2023) To enhance network security and mitigate the risks posed by cyberattacks, there is a growing imperative to leverage advanced technologies such as artificial intelligence (AI) to develop automated incident response systems. (Maddireddy and Maddireddy, 2023) Using AI algorithms and machine learning to enhance issue detection and response in cloud environments became increasingly important in 2020 (Reddy & Ayyadapu 2020). AI-driven responses dramatically shorten the window between detection and mitigation, drastically reducing the potential impact of cyber incidents. This shift towards automation and predictive analysis heralds a new era of cybersecurity strategy, where responses are not only faster but also more intelligently focused on the most critical threats. (Wilson, 2019). Thus, integrating AI into the cyber incident response is a critical step in creating a more reliable, automated, and efficient defense

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VIII August 2025

system against the constantly changing array of cyber threats. These systems can autonomously detect anomalies, correlate security events, and orchestrate response actions in a fraction of the time it would take for human intervention, thereby reducing the dwell time of attackers and minimizing the potential impact of security breaches. (Maddireddy and Maddireddy, 2023). This change is represented by the AI Cyber Incident Response System project, which aims to transform how businesses respond to cyber incidents by making the procedure more intelligent, efficient, and ultimately secure. This research work aims to develop an AI-assisted detection system that not only detects cyber threats with unprecedented accuracy but also categorizes them based on severity, this research addresses a critical gap in current cybersecurity practices. The introduction of automated response protocols further enhances the ability of organizations to respond to cyber threats swiftly and effectively, minimizing the impact of attacks and reducing operational downtime. This integrated approach represents a significant advancement in the field of cyber security and provides a strong framework for protecting digital assets in the digital age. It promises a more dynamic, efficient, and effective defense against the increasingly sophisticated cyber threats that organizations face today.

Related Work

Jatin and Neha (2023) explored the critical role of Artificial Intelligence (AI) in enhancing cybersecurity strategies in their study. Their research sheds light on the diverse applications of AI technologies in cybersecurity, particularly emphasizing the use of machine learning and analytics for real-time threat detection and response. This utilization of AI enables the swift and accurate analysis of large datasets to identify and act upon anomalies indicative of cyber threats, marking a significant advancement in cybersecurity operations. Pahuja et al., (2023) also contemplate the future of cybersecurity, positing that AI's potential to predict and prevent threats could lead to substantially more effective and resilient security measures. This foresight and AI's automation capabilities may significantly narrow the opportunities for cyberattacks, minimizing their potential damage. Moreover, integrating AI into cybersecurity frameworks offers prospects for morestreamlined and cost-efficient security operations. The study not only acknowledges the current advantages of AI in fortifying cybersecurity but also anticipates future developments that could further solidify digital defenses against an array of cyber threats. Alturkistani et al., (2022) delve into the transformative impact of deep reinforcement learning (DRL) on cybersecurity, specifically in enhancing post-alert decisions within Security Incident and Event Management (SIEM) systems. The research makes a compelling argument for integrating DRL into SIEM systems, underscoring its ability to make accurate, real-time decisions independent of preexisting datasets. The authors provided an innovative framework designed to seamlessly integrate Incident Response Plans (IRPs) with the APIs of various security tools. This framework addresses the complex challenge faced by Security Operation Centers (SOC) of manually correlating IRPs with corresponding APIs, streamlining what has traditionally been a cumbersome and time-intensive process. Kaiser et al., (2022) presented a transformative approach in the field of cybersecurity incident response, marked by the introduction of the 'Attack Incident Responder,' a method predicated on simple heuristics designed to autonomously formulate countermeasures against cyber threats. A pivotal element of the research is the exploration of the integration between automated incident response and proactive threat hunting, a combination that propels the cybersecurity domain toward the ideal of fully automated Security Operation Centers (SOC). This integration facilitates a dynamic cybersecurity environment where continuous threat monitoring is paired with immediate response actions, substantially altering the operational landscape of cybersecurity. The research further introduces a novel precision threshold concept for generating attack hypotheses, a critical metric for assessing the performance of predictive defense algorithms against conventional defense mechanisms.Ban et al., (2023) integrate AI with cybersecurity, enhancing Security Information and Event Management (SIEM) systems to combat low-quality IDS notifications and alert fatigue through automation and response capabilities. The framework uses advanced machine learning algorithms and data visualization to reduce false alarms and address alert fatigue in cybersecurity operations. It integrates host-based and network-based IDS, enhancing operational capabilities. The XGBoost-MAS-OSS algorithm and Weighted Support Vector Machine method improve performance, while the WSVM method reduces false positive rates and achieves the highest recall rate. The study by Ban et al., (2023) highlights the significant impact of AI- enhanced SIEM systems on operational efficiency, highlighting a shift towards AI and machine learning to tackle cybersecurity challenges.

Vast et al., (2021) developed an advanced Security Orchestration, Automation, and Response (SOAR) system to address the complexities of cyber threats in the digital age. The system uses deep-learning techniques to collect

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VIII August 2025

data from sources like firewalls and IDS, improving detection accuracy. It integrates with SIEM systems to produce comprehensive reports, strengthening detection and response effectiveness. The system's precise assessment of threat severity aids in swift decision- making in combating cyber threats. Vast et al.'s research advocates for intelligent, automated, and adaptive cybersecurity frameworks. They highlight the transformative potential of AI in redefining digital infrastructure protection against cyber threats

The study emphasizes the need for predictive security systems, anticipating threats and automating responses in real-time. This shift is crucial in today's connected world, where the proliferation of digital devices outpaces traditional security protocols.

Correa et al., (2021) introduce a transformative autonomic computing architecture to improve Cyber-Security Incident Response Team Intelligent Decision Support Systems (CSIRT- IDSS) by prioritizing and suggesting immediate mitigation strategies. Correa's approach to cyberattack countermeasures involves integrating Web Service Oriented LP, Constraint LP, and Object-Oriented LP to create a COP to combat various cyberattack vectors. The model's effectiveness is demonstrated through CARMAS, a prototype tested against simulated cyberattacks. The author proposes a multi-objective COP, considering factors like service availability and mitigation strategy costs, which could revolutionize decision-making within CSIRT-IDSS. The research demonstrated the use of logic programming techniques to improve cybersecurity incident response (CSIRT) effectiveness, highlighting the need for adaptable approaches in the evolving digital threat landscape.. Pujar et al., (2021) proposed a model that identifies, detects, categorizes, and responds to attacks, generating warnings and alarms. Implementing this model ensures better protection of wireless network systems. The model comprises of Detection Phase, Response System, and Security Mechanisms. The detection system and response system will be integrated to automate responses to detected incidents. Using anomaly and signature detection in a static-based response system to identify suspicious activity in wireless networks. The authors created and developed a model to respond to events in an efficient manner, choosing response measures to guarantee improved security of wireless network systems using the AES encryption technique.

Shah, V. (2021) discusses the role of machine learning algorithms in cybersecurity, focusing on their ability to detect and prevent various threats. Machine learning algorithms use data-driven techniques to analyze vast amounts of information, identifying patterns and anomalies indicative of malicious activities. They can detect hidden threats, evade traditional signature-based detection methods, and analyze unstructured data types like network traffic and user behavior. In threat prevention, machine learning algorithms play a crucial role in proactive defense strategies, allowing organizations to implement preemptive measures before attacks occur.

The multifaceted nature of cyber incidents makes it difficult to determine the success of incident response efforts. However, by establishing metrics and evaluation frameworks to measure and improve incident response performance, cyber incident response systems can be considerably more resilient and effective. This will help organizations better defend themselves against continuously developing cyber threats

METHODOLOGY

This research outlined five critical stages: data engineering, model training for incident detection, model testing and evaluation, the development of a severity-based incident arrangement system, and the implementation of automated response steps as shown in the system architecture in Figure 1. Understanding the system's organization and designing its overall structure are key components of architectural design.

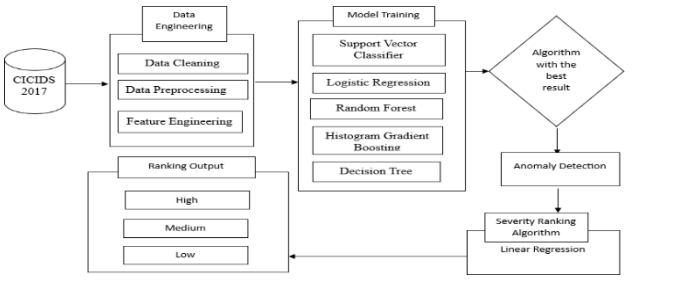


Figure 1: Architecture of the Cyber Incident Response Systems

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VIII August 2025

Data Collection

The dataset used for this research work is the CICIDS2017 dataset obtained from Kaggle. It is an academic intrusion detection dataset. It was published by the Canadian Institute for Cybersecurity. The benign and up-to-date attacks that simulate realistic data from the real world are included in the CICIDS 2017 dataset. It also contains the outcomes of the network traffic analysis, with labelled flows according to the time stamp, source and destination IP addresses, source and destination ports, protocols, and attack types.

Data Engineering Data preprocessing is essential for maintaining the integrity and reliability of the model's output, ensuring that it learns from accurate and relevant data. A selection was made of the most relevant features that enhance the model's predictive ability. The data preparation process involves the removal of any data anomalies.

Model Training and Evaluation: The first phase in model training is the model selection, where a range of machine learning algorithms are considered, including the Logistic Regression, Random Forest, Support Vector Classifier, Decision Tree Classifier, and Histogram Gradient Boosting Classifier. The comparative analysis is pivotal, as it not only sheds light on the performance dynamics of each model but also guides the selection process, ensuring that the most suitable algorithm is advanced for further refinement.

Through a cycle of testing, feedback, and subsequent optimization, the models are fine-tuned to enhance their threat detection capabilities. This iterative process is crucial for minimizing false positives and negatives, thereby increasing the reliability and accuracy of the incident detection mechanism.

This phase ends when an algorithm is found that not only satisfies but also surpasses the requirements for successful cyber threat detection. The selected algorithm, having been vetted and optimized, lays the groundwork for a system that can accurately and swiftly identify cyber threats. After a comprehensive evaluation, the best performing model based on its performance in metrics like accuracy, precision, and F1-score was picked and input into the Severity ranking Algorithm that determines the extent of the attack detected. This capability is paramount, as it directly influences the system's ability to facilitate timely and effective incident response, a critical component in mitigating the impact of cyberattacks. Moreover, this phase underscores the project's commitment to leveraging cutting-edge AI technologies to bolster cybersecurity defences, setting a new benchmark for excellence in cyber incident response.

The evaluation process commences with the calculation of key performance metrics, which serve as quantifiable indicators of each model's effectiveness. Accuracy (ACC), a fundamental metric, is calculated using the formula

$$CC = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)$$

Where TP (True Positives) and TN (True Negatives) represent correctly identified threat events and non-threat events, respectively, while FP (False Positives) and FN (False Negatives) denote erroneously identified events. Precision as positive predictive value (Cox & Vladescu,2023)

$$Precision = \frac{TP}{TP + FP}, \quad (2)$$

Sensitivity or recall is also known as the true positive rate (Dehmer & Basak, 2012)

$$Recall = \frac{TP}{TP + FN}. \quad (3)$$

The F1 Score, which harmonizes the balance between Precision and Recall,]

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \quad (4)$$

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VIII August 2025

offering a single metric to assess the model's overall performance, Squared Test, articulated through the formula

where O_i and E_i are the observed and expected frequencies, is instrumental in evaluating the disparity in performance across models. ANOVA further complements this analysis by employing the F-statistic,

$$F = \frac{Between-group\ variability}{Within-group\ variability} \quad (6)$$

to verify if the performance means of different models are statistically distinct, providing a robust framework for comparative evaluation. This detailed approach to model evaluation underscores the commitment to deploying a system that is not only grounded in statistical excellence but also primed for efficacy in the multifaceted domain of cybersecurity

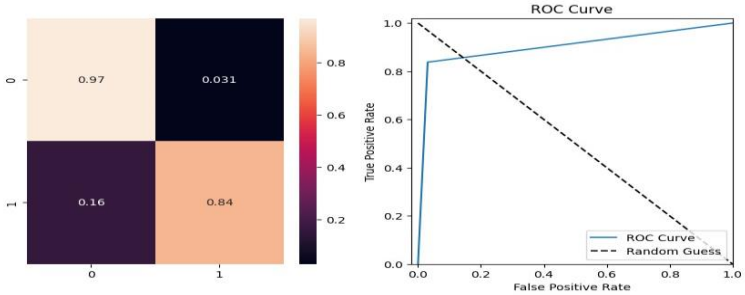
Severity-Based Incident Arrangement System

This progression is characterized by the adoption of an algorithmic approach that harnesses the intrinsic characteristics of cyber incidents to autonomously determine their severity. It is designed to quantify the threat level of each incident through a precise mathematical formula, thereby streamlining the prioritization process in incident response strategies

$(i) = \alpha \cdot I(i) + \beta \cdot T(i) + \gamma \cdot E(i)$ (7) This function is designed to compute the severity score $S(i)$ of of a given cyber incident i. This computation incorporates three fundamental parameters: the intensity (i), frequency $T(i)$, and potential extent of impact $E(i)$ of the incident.

Automated Incident Response Protocol Development

This procedure includes developing a variety of response protocols, each intended to counter particular categories of cyberthreats, from mild anomalies to serious, targeted assaults.



The development of these protocols is deeply rooted in a thorough understanding of the threat landscape, leveraging the insights provided by incident detection systems to ensure that responses are precisely calibrated to the nature and severity of the threat. The seamless integration of these automated response protocols with AI-driven threat detection systems is a cornerstone of modern cybersecurity strategy. Figure 2 illustrates the sequence diagram that displays the process flow of a cyber incident response in an organization using an AI-enhanced system and a Response Team

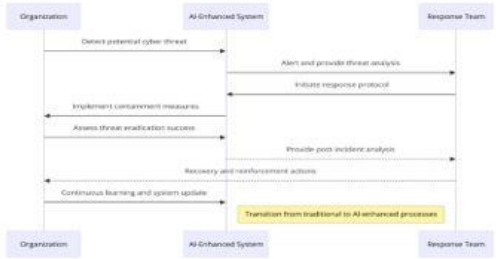


Figure 2 A sequence diagram on Cyber Incident Response

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VIII August 2025

Implementation

The efficacy of an IDS relies heavily on the choice of models and their configuration. Each model's performance is evaluated using the test data set aside during preprocessing.

Logistic Regression

The Logistic Regression model is a foundational machine learning algorithm known for its simplicity and efficiency in binary classification tasks. In the context of IDS, it works by estimating the probabilities of network events being benign or attacks based on the logistic function. The performance metrics for the Logistic Regression model reveal a strong ability to differentiate between the classes. The precision metric for the benign class (labeled as '0') is 0.97, indicating that the model has a high accuracy rate when predicting normal traffic. The recall for the same class is also 0.97, which means the model is equally adept at identifying true benign instances. Conversely, for the attack class (labeled as '1'), the precision is lower at 0.85, with a recall of 0.84, suggesting that while still performing well, the model is somewhat less reliable in classifying attack instances.

Overall, the model achieves an accuracy of approximately 94.7%, a very respectable score that underscores its effectiveness. However, when it comes to the balance between precision and recall, the F1 Score, which is the harmonic mean of the two, is a critical measure, particularly in domains like cybersecurity where the cost of false negatives (failing to detect an attack) and false positives (incorrectly blocking benign activity) can be high. For the attack class, the F1 Score is approximately 0.84, which indicates good but not excellent performance.

The ROC curve in Figure 3 further complements these findings. It shows the trade-off between the true positive rate (sensitivity) and the false positive rate (1 - specificity) at various threshold settings. The curve is significantly above the line of randomness (dashed line), which is indicative of the model's good classification ability. The closer the curve follows the left-hand border and then the top border of the ROC space, the more accurate the test. In summary, the Logistic Regression model shows a strong capability to classify benign traffic accurately while offering a reasonable degree of reliability in identifying attacks within the CICIDS2017 dataset. Despite its simplicity, it provides a robust baseline for IDS performance, and its statistical outputs, particularly the ROC curve, provide valuable insights for further refinement and comparison with more complex models.

Random Forest Classifier

The Random Forest model, known for its robustness and accuracy in various machine learning tasks, appears to have delivered exceptional results. The model's performance metrics suggest near-perfect precision, recall, and F1 scores for both benign and attack classes, indicating an almost ideal classifier in this context. With precision and recall scores at the threshold of

1.00 for both classes, the model demonstrates a high level of accuracy in correctly identifying true positives and true negatives. This performance is remarkable and suggests that the Random Forest model is exceptionally well-tuned.The accuracy metric of approximately 99.9% further confirms the model's capability to classify network traffic with minimal error.

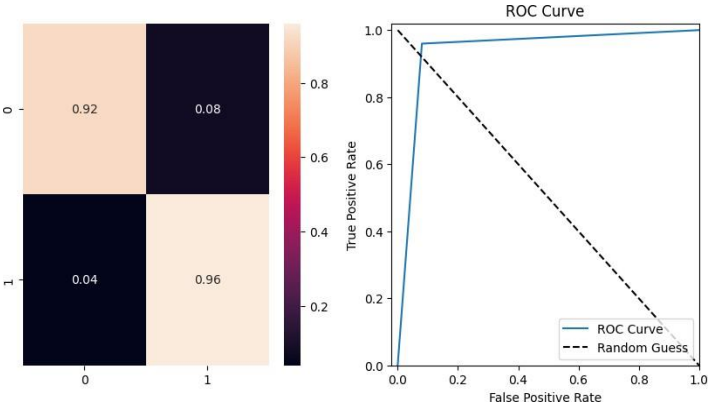


Figure 4: Heatmap and ROC curve for Random Forest Classifier

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VIII August 2025

The heatmap in Figure 4 shows the confusion matrix, which provides a visual representation of the model's classification accuracy. The values along the diagonal, which are nearly 1, represent a high number of correct predictions for both benign

(0) and attack (1) classes, with negligible false positives and false negatives, as shown by the near-zero values off the diagonal. The area under the curve (AUC) is nearly 1, indicating that the model has an excellent measure of separability and can distinguish between the two classes with high confidence. In conclusion, the Random Forest model has been shown to be a highly effective algorithm for the IDS task.

Histogram Gradient Boosting Classifier

The Histogram Gradient Boosting Classifier, a machine learning algorithm known for its effectiveness in handling large datasets and its efficient use of histogram-based optimization for gradient boosting, also showcases an outstanding performance.. The provided metrics point towards an exceptional ability to classify the traffic correctly, with precision and recall rates at or very close to 1.00 for both classes.

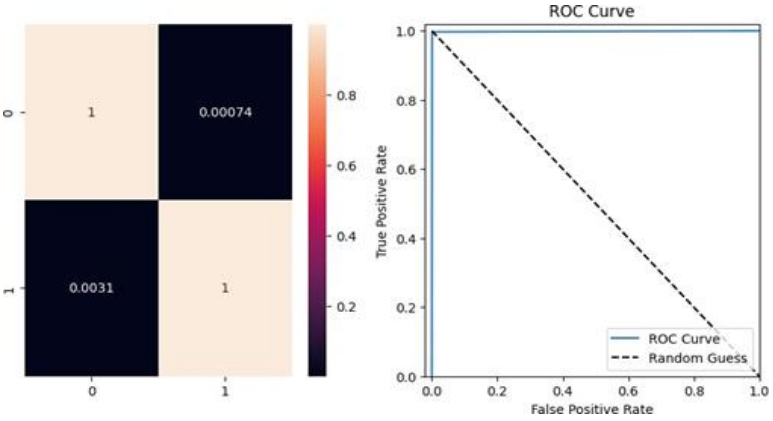


Figure 5: Heatmap and ROC Curve for Histogram Gradient Boosting Classifier

The confusion matrix depicted in the heatmap in Figure 5 is almost perfectly diagonal, indicating that the Histogram Gradient Boosting Classifier has correctly identified nearly all benign and attack instances. The minimal values of the diagonal show that there are very few false positives and false negatives, highlighting the model's accuracy. Complementing the confusion matrix, the ROC Curve indicates excellent model performance, with a high area under the curve (AUC), suggesting that the model can distinguish between benign and attack classes with high confidence. In summary, the Histogram Gradient Boosting Classifier seems to be highly effective for this IDS task, demonstrating excellent classification ability.

Linear Support Vector Classifier

The Linear Support Vector Classifier (Linear SVC) is a variation of the Support Vector Machine (SVM) that is particularly well-suited for classifying high-dimensional data. The performance metrics for the Linear SVC applied suggest a high degree of effectiveness in distinguishing between benign and attack instances, though not as high as the previously discussed models.

The precision for the benign class is 0.96, and the recall is 0.98, indicating that the model is very capable of identifying benign traffic accurately. However, for the attack class, while the precision is still relatively high at 0.88, the recall is lower at 0.82. This lower recall rate implies that the Linear SVC is less effective at identifying all actual attack instances, as a recall rate below 1.00 indicates that there are more false negatives, where attacks are incorrectly labeled as benign.The heatmap shows the confusion matrix, where the values on the diagonal represent correctly classified instances. While the majority of benign cases are correctly identified (as indicated by the 0.98 value), the model struggles with attack instances, missing 18% of them, as evidenced by the 0.18 value in the false negative area. The Receiver Operating Characteristic (ROC) is significantly above the line representing random guessing, which indicates a good predictive performance. However, it does not get as close to the top-left corner as the curves for the previous models, reflecting the lower recall rate. In terms of overall accuracy, the model achieves 95.02%, which is still a strong performance but indicates that there is room for

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VIII August 2025

improvement, especially in correctly identifying attacks. The macro-average F1-score of

0.91 suggests that the model is generally balanced between precision and recall across both classes .In conclusion, the Linear SVC demonstrates good performance for the IDS task on the CICIDS2017 dataset, but may require further tuning or ensemble methods to improve its detection rate of attack instances. The results emphasize the importance of considering multiple metrics beyond accuracy when evaluating models for imbalanced classification tasks like IDS.

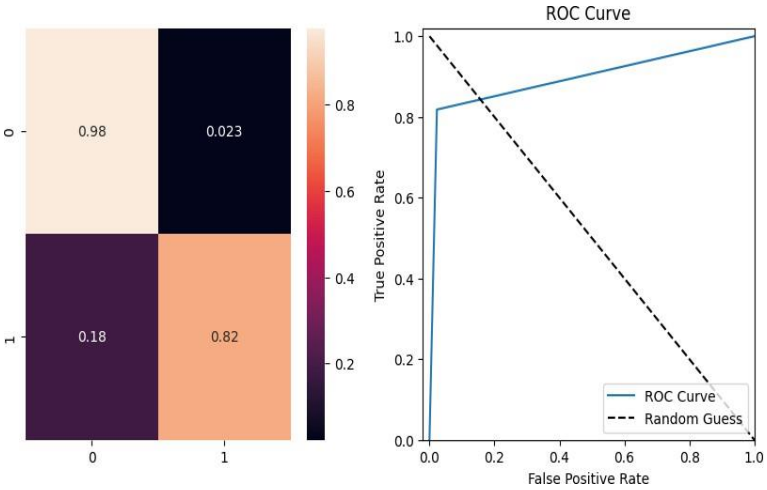


Figure 6: Heatmap and ROC curve for Linear Support Vector Classifier

Decision Tree Classifier

The Decision Tree Classifier is a widely used method that creates a model in the form of a tree structure, making decisions by splitting data based on certain criteria. The model's evaluation on the CICIDS2017 dataset indicates very high performance across all metrics, with an overall accuracy of nearly 99.85%. The precision score of 99.55% and the recall of 99.61% for the Decision Tree model suggest that it has an excellent capability to distinguish between benign and attack instances. These scores are reflected in the confusion matrix, the false positives and false negatives are extremely rare, as indicated by the very low values outside the diagonal. In fact, the model has nearly perfect recall for the benign class and very high recall for the attack class.

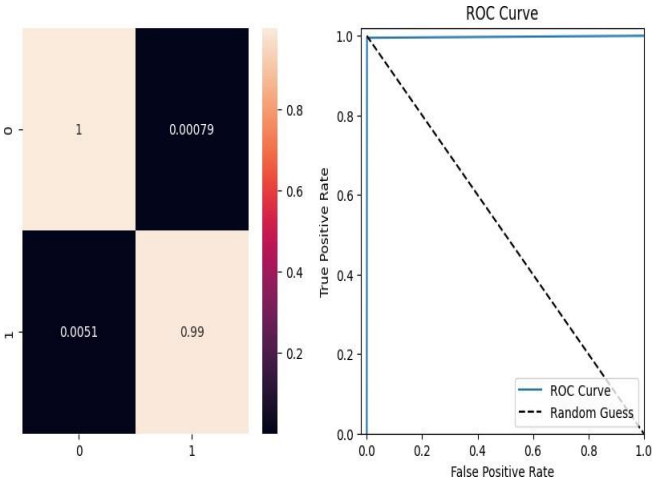


Figure 7: Heatmap and ROC curve for Decision Tree Classifier

The ROC Curve shows that the Decision Tree Classifier performs significantly better than random guessing, which is represented by the dashed diagonal line. The curve approaches the top left corner of the plot, indicative of a high true positive rate and a low false positive rate across different thresholds.

While these results are commendable, it's important to note that decision trees can be prone to overfitting, especially if they are deep with many branches. This can make them sensitive to the specific noise in the training

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VIII August 2025

set, which may lead to excellent training performance but poor generalization to new data. Therefore, these results should be validated with additional testing, such as cross-validation or using a different test set, to ensure that the model generalizes well and that its performance is not an artifact of overfitting. In summary, the Decision Tree Classifier appears to perform exceptionally well on the dataset, with scores suggesting that it can effectively separate benign traffic from cyber-attacks. However, caution should be taken to confirm these results through further validation techniques to ensure that the model's predictive power holds true in practice. The table below displays the results from the machine-learning algorithm used above.

Table 1: A Comparison of the Results of Selected Machine Learning Algorithms

Model	Accuracy	F1-Score	Precision
Logistic Regression	94.7%	0.84	0.84
Random Forest Classifier	99.9%	0.99	0.99
Histogram Gradient Boosting	99.9%	0.99	0.99
Support Vector Classifier	95.0%	0.82	0.84
Decision Tree Classifier	99.8	0.99	0.99

Severity Score Ranking Algorithm

In the Severity Ranking Algorithm, the severity (i) of an incident i is computed by considering three fundamental attributes of the incident: intensity $I(i)$, frequency $T(i)$, and the potential extent of damage $E(i)$. These attributes are synthesized into a single severity score through a weighted sum, where α, β, and γ serve as the respective weights. To apply this algorithm to the CICIDS2017 dataset, one would start by calculating the intensity (i) for each incident, which reflects the strength or magnitude of the incident at the moment of its occurrence. Intensity could be derived from various indicators such as the size of the attack payload, the number of systems compromised, or the level of network activity disruption caused. The frequency

(i), which pertains to how often an incident occurs, would be determined. This could involve analyzing the dataset for repeated patterns or rates of particular attack types. An attack that occurs with greater frequency may suggest a persistent threat requiring more urgent attention. Finally, the potential extent of damage, or potential damage (i), is assessed. This might consider the possible consequences of an attack, including data loss, service interruption, or other forms of system impairment. The potential damage is a forward-looking metric, projecting the possible outcomes if the incident is not appropriately mitigated. By integrating these computed values into the dataset, each incident is assigned a severity score that encapsulates the overall threat it represents, considering not just the occurrence of an attack but also its implications. This score is instrumental in prioritizing incidents for response, allocating resources more efficiently, and bolstering the overall security posture. The calculated severity scores could also enhance the predictive power of machine learning models applied to the dataset by providing a rich, quantitative measure of each incident's seriousness.

The intensity of the Attack

Calculating the intensity of an incident within the CICIDS2017 dataset involves aggregation of various network traffic features that collectively represent the vigor or forcefulness of the incident. The intensity is conceived as a composite measure reflecting the volume and the characteristics of the traffic associated with the incident. It encapsulates the sum of packets transmitted in both directions—forward and backward— indicating the level of communication between hosts. Additionally, the mean lengths of the packets moving forward and backward are considered, providing a sense of the size of the data being exchanged, which can be indicative of the load or the potential payload of an attack. Further contributing to the intensity are several flag counts that are crucial in the TCP/IP protocol, each representing different control mechanisms within the network communication. These include flags for pushing data, synchronization of sequence numbers, reset signals, acknowledgment of the

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue VIII August 2025

receipt of packets, urgent pointers indicating the presence of priority data, and various control flags used for window scaling, congestion notification, and other purposes. The presence and frequency of these flags in network traffic can signal unusual or potentially malicious activity, such as repeated attempts to establish a connection or the urgent transmission of data, which might suggest an ongoing attack or exploitation attempt.

By summing these values, a single intensity score is obtained for each incident, offering a quantified reflection of its network traffic profile. This score is valuable for understanding the immediate impact of the incident, gauging its potential threat level, and for subsequent steps in the security analysis, such as prioritization of incidents or informing the development of defensive strategies. The composite nature of this intensity measure helps in distinguishing between benign network behavior and potentially harmful traffic that could warrant a deeper investigation or immediate action.

Frequency of the Attack

The frequency of an incident within the CICIDS2017 dataset is assessed by focusing on two dynamic attributes of the network traffic: the rate of packets per second and the rate of bytes per second. These two metrics serve as proxies for the regularity and repetition of network activities associated with an incident. The packets per second rate offer a direct measure of how many packets are being sent through the network over time, which can vary from normal to extraordinarily high during an attack, such as a denial-of-service incident. A higher frequency of packets could indicate a more aggressive attack or a higher level of network activity associated with an incident. Similarly, the flow of bytes per second reflects the volume of data transmitted across the network. An increase in this rate might be symptomatic of large data transfers, which could be ordinary in a data-intensive operation but might also signify data exfiltration or a flooding type of attack when present in a security incident context. The frequency measure is constructed to represent the overall tempo of the network traffic, encapsulating both the number and the size of the network packets in the flow. This integrated perspective on frequency provides an ongoing assessment of the network condition, helping to identify and prioritize incidents that manifest as vital indicator for cybersecurity defences, as frequent and high- volume network traffic can be a harbinger of severe security threats that may require immediate attention.

Potential Extent of Damage or Potential Impact

In evaluating the potential extent of damage or impact from a cybersecurity incident within the CICIDS2017 dataset, the calculation integrates several parameters that collectively offer an estimate of the possible consequences of an incident. The destination port number is considered as part of this calculation, reflecting the target of the network traffic. Certain ports are commonly associated with specific services or applications, some of which may be critical and, if compromised, could lead to significant disruption or damage. The duration of the flow is another critical factor, providing insight into how long the network traffic associated with the incident persisted. Longer durations could imply a sustained attack, increasing the likelihood of successful penetration or data loss, thus potentially exacerbating the impact. The average bytes per bulk in the forward and backward directions indicate the volume of data involved in bulk transfer events during the incident. High volumes might suggest substantial data movement, which in the context of an incident could mean data theft or the transfer of a large volume of malicious data into the network. By aggregating these factors, the dataset quantifies an incident’s impact extent, giving a multifaceted view of its potential to inflict damage or cause substantial harm. This measure helps to prioritize incidents based on the severity of their possible impact, guiding the response strategy to focus resources where they are most needed to mitigate risk and minimize damage.

Alpha, Beta, and Gamma Coefficients

To determine the appropriate weights for the intensity, frequency, and potential impact in the Severity Ranking Algorithm, a linear regression model is employed using the specific features of incidents as the predictors and the types of attacks as the response variable. It is crucial to understand that the dataset has been refined to focus solely on attack instances, with benign samples removed, ensuring that the model's coefficients reflect the characteristics of the attacks alone. By training the linear regression model on this attack-only dataset, the algorithm learns to associate the varying levels of intensity, frequency, and impact with the different types of attacks. The model's resulting coefficients—obtained from the 'model.coef_' attribute post-training—serve as a

quantified measure of the contribution of each feature towards the classification of an attack.

In essence, the coefficient α associated with intensity reveals how strongly the intensity of an attack correlates with its classification. A higher value would suggest that more intense attacks are likely to be classified differently than less intense ones. Similarly, the coefficient β related to frequency sheds light on the importance of the occurrence rate of attacks, and γ for the potential impact offers insights into how the possible consequences of an attack influence its classification.

The values of α, β, and γ distilled through the linear regression model, encapsulates the relative significance of each predictor in determining the severity of an attack. This technique leverages the predictive capability of linear regression to translate complex relationships within the data into a simple yet powerful set of weights, which can then be used in the Severity Ranking Algorithm to score and prioritize incidents based on the calculated severity scores.

Severity Threshold Marking

Upon obtaining the coefficients from the linear regression model and calculating the severity scores for each incident, the next step is to determine the range of severity within the dataset. This is done by identifying the maximum and minimum severity scores. With these values, it is possible to establish thresholds that categorize incidents into different levels of severity. The range of severity is divided into three equal segments, representing low, medium, and high severity levels. The lowest point of the severity score range is taken as the starting point, and the range is divided to determine the points that separate the low from the medium, and the medium from the high-severity incidents. These thresholds are critical for incident response prioritization, as they enable security analysts to identify which incidents require immediate attention, which are of moderate urgency, and which can be attended to routinely. This stratification ensures that resources are allocated efficiently, focusing efforts where they are most needed to protect against cyber threats. It provides a systematic approach to security management, enhancing the effectiveness of the incident response by quantifying the severity of each detected incident.

CONCLUSION

The development and evaluation of the Cyber Incident Response System (CIRS) have demonstrated the significant potential of machine learning in enhancing cybersecurity defences. The journey from system specification to the thorough analysis of the CICIDS2017 dataset has highlighted the critical role of data- driven approaches in modern cybersecurity practices. The application of models such as Logistic Regression, Random Forest, and Histogram Gradient Boosting Classifiers yielded impressive accuracy in detecting and classifying cyber threats. The Severity Ranking Algorithm represented a key advancement, allowing for a more precise assessment of incident severity. By quantifying the intensity, frequency, and potential impact of threats, the algorithm facilitated more effective prioritization of responses, ensuring that critical incidents received the attention they required

REFERENCE

1. Jatin Pahuja and Neha Agrawal (2023). AI in cyber security. International journal of communication and information technology, 4(1):46-53. doi: 10.33545/2707661x.2023.v4.i1a.59
2. Cranford, J. (2023, July 7). Incident Response Plan: Frameworks and Steps. CrowdStrike. https://www.crowdstrike.com/cybersecurity-101/incident- response/incident-responsesteps/
3. Wilson, S. (2019). AI security tech is making waves in incident response. TechTarget. https://www.techtarget.com/searchcio/feature/AI-security-tech- is-making-waves-inincident-response.
4. Alturkistani, H., & El-Affendi, M. A. (2022). Optimizing cybersecurity incident response decisions using deep reinforcement learning. International Journal of Electrical and Computer Engineering, 12(6), 6768. DOI: 10.11591/ijece.v12i6.pp6768-6776
5. Kaiser, F. K., Andris, L. J., Tennig, T. F., Iser, J. M., Wiens, M., & Schultmann, F. (2022, October). Cyber threat intelligence enabled automated attack incident response. In 2022 3rd International Conference on Next Generation Computing Applications (Next Comp) (pp.1-6). IEEE. **DOI:**

10.1109/NextComp55567.2022.9932254

6. Ban, T., Takahashi, T., Ndichu, S., & Inoue, D. (2023). Breaking alert fatigue: Ai-assisted siem framework for effective incident response. Applied Sciences,13(11), 6610.**https://doi.org/10.3390/app13116610**
7. Vast, R., Sawant, S., Thorbole, A., & Badgujar, V. (2021, April). Artificial intelligence based security orchestration, automation and response system. In 2021 6th International Conference for Convergence in Technology (I2CT) (pp. 1-5). IEEE. DOI: 10.1109/I2CT51068.2021.9418109
8. Pierazzi, F., Casolari, S., Colajanni, M., & Marchetti, M. (2016). Exploratory security analytics for anomaly detection. Computers & security, 56, 28 49. https://doi.org/10.1016/j.cose.2015.10.003
9. Naseer, A., Naseer, H., Ahmad, A., Maynard, S. B., & Siddiqui, M. (2021). Real-time analytics, incident response process agility and enterprise cybersecurity performance: A contingent resource-based analysis. International Journal of Information Management, 59, 102334.https://doi.org/10.1016/j.ijinfomgt.2021.102334
10. Pujar, S. N., Choudhary, G., Shandilya, S. K., Sihag, V., & Choudhary, A. (2021). An adaptive auto incident response based security framework for wireless network systems. Research Briefs on Information & Communication Technology Evolution, 7, 4 DOI 10.22667/ReBiCTE.2021.08.15.004
11. Cox, D. J., & Vladescu, J. C. (2023). Statistics for applied behavior analysis practitioners and researchers. Elsevier.
12. Dehmer, M. & Basak, S. C. Statistical and Machine Learning Approaches for Network Analysis DOI:10.1002/9781118346990
13. Staves, A., Anderson, T., Balderstone, H., Green, B., Gouglidis, A., & Hutchison, D. (2022). A cyber incident response and recovery framework to support operators of industrial control systems. International Journal of Critical Infrastructure Protection,37, https://doi.org/10.1016/j.ijcip.2021.100505
14. Maddireddy, B. R., & Maddireddy, B. R. (2023). Enhancing Network Security through AI-Powered Automated Incident Response Systems. International Journal of Advanced Engineering Technologies and Innovations, 1(02), 282-304.
15. Correa, C., Robin, J., Mazo, R., & Abreu, S. (2021). Intelligent decision support for cybersecurity incident response teams: autonomic architecture and mitigation search. In International Conference on Risks and Security of Internet and Systems (pp. 91-107). Cham: Springer International Publish