IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 29, No. 6, December, 2024, pp. 2415-2424
Journal homepage: http://iieta.org/journals/isi

Machine Learning-Based Anomaly Detection Model for Cybersecurity Threat Detection

Ubaid Mohamed Dahir[1*], Abdirahman Osman Hashi[1], Abdullahi Ahmed Abdirahman[1], Mohamed Abdirahman Elmi[1], Octavio Ernest Romo Rodriguez[2]

[1] Faculty of Computing, SIMAD University, Mogadishu 252, Somalia
[2] Department of Computer Science, Faculty of Informatics, Istanbul Teknik Universitesi, Istanbul 34467, Turkey

Corresponding Author Email: engubaid@simad.edu.so

https://doi.org/10.18280/isi.290628

Received: 25 April 2024
Revised: 21 September 2024
Accepted: 28 September 2024
Available online: 25 December 2024

Keywords:
cybersecurity, anomaly detection, machine learning, naive bayes classifier, multinomial model, threat classification

ABSTRACT

The proliferation of cybersecurity threats continues to challenge the resilience of information systems worldwide. An effective defense against such threats requires advanced detection methods that can predict and classify the severity of vulnerabilities with high precision. This paper proposes a sophisticated anomaly detection framework using a machine learning algorithm, aimed at identifying and categorizing cybersecurity vulnerabilities from the CISA Known Exploited Vulnerabilities catalog for 2022. The proposed model underwent a rigorous process of preprocessing and data cleaning to ensure the integrity and suitability of the data for machine learning analysis. It has demonstrated exceptional proficiency, achieving an accuracy rate of 0.9810, alongside high precision and recall values across various severity levels of vulnerabilities. The model's performance highlights its utility in enhancing cybersecurity measures. Therefore, the significance of this model lies in its potential to transform the field of cybersecurity, offering a scalable, efficient tool for proactive threat detection and contributing to the fortification of information systems against a broad spectrum of cyber threats.

1. INTRODUCTION

In today's digital age, cybersecurity has become paramount. Organizations across the globe increasingly rely on digital infrastructures, and consequently, the potential impacts of cyber threats continue to grow. Cybersecurity is crucial not only for protecting information systems from unauthorized access and attacks but also for maintaining the functionality and trust that are foundational to modern organizations [1]. The importance of cybersecurity cannot be overstated, as its role extends beyond mere protection of data to safeguarding the very integrity of our digital way of life. As cyber threats have evolved, so too have the methods to combat them. Initially considered mere nuisances, cyber threats have progressively become more sophisticated, aimed at stealing data, extorting money, or disrupting services. This evolution from simple to complex threats has necessitated a corresponding shift in defence strategies. Traditional security measures often rely on establishing strong perimeters and defending known vulnerabilities with predefined rules. However, these methods are increasingly insufficient, primarily because they struggle to adapt to novel or evolving threats, often leading to significant security breaches [2].

The concept of anomaly detection has thus come to the fore as a critical component in cybersecurity strategies. Anomaly detection involves identifying patterns in data that do not conform to expected behaviour. In the context of cybersecurity, this means spotting potential threats that deviate from normal network or system activities [3]. The dynamic nature of cyber threats makes them particularly challenging to predict and manage, underscoring the importance of detecting anomalous behaviour as a clue to potential breaches. Machine learning has emerged as a powerful tool in enhancing the capabilities of anomaly detection systems. By leveraging historical data, machine learning models can detect subtle and complex patterns indicative of malicious activity. These models offer a significant improvement over traditional methods, particularly in their ability to adapt to new and unforeseen attack vectors. Despite these advantages, anomaly detection in cybersecurity is not without challenges, such as high false-positive rates and the need for large, labelled datasets [4].

The schoolers have responded to these challenges by proposing various innovative machine learning models to enhance the accuracy and efficiency of anomaly detection systems. For instance, Hernandez-Jaimes et al. [5] have explored the use of neural networks in anomaly-based intrusion detection, demonstrating a marked improvement in detection capabilities. Moreover, the advent of deep learning—a subset of machine learning that utilizes complex neural network architectures—has further advanced the field. Deep learning techniques, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have proven particularly effective in identifying intricate patterns that are typical of modern cyber-attacks. A compelling example of these advancements is the work by Delchevalerie et al. [6] who applied CNNs to detect Distributed Denial of Service (DDoS) attacks, one of the most prevalent threats in network security. Their findings illustrated not only improved accuracy but also increased speed in detecting such attacks compared to traditional methods. However, integrating these sophisticated machine learning models into existing cybersecurity systems presents its own set of challenges, including computational demands and the need for real-time processing capabilities.

Meanwhile, Unsupervised learning models, which do not require labelled data, offer promising solutions in scenarios where threat labels are scarce. Techniques such as clustering and autoencoders are being explored for their potential to autonomously detect unknown types of cyber threats [7]. Furthermore, hybrid models that combine various machine learning approaches can potentially lead to more robust anomaly detection systems by leveraging the strengths of both supervised and unsupervised learning methods. The practical implications of implementing these advanced machine learning models are profound. They have the potential to significantly reduce the time required to detect and respond to threats, minimize the occurrence of false positives, and optimize the allocation of human security resources. Looking forward, the future of anomaly detection in cybersecurity will likely focus on enhancing the adaptability and scalability of these models to keep pace with the continuously evolving landscape of cyber threats [8].

This research examines how the use of anomaly detection methods, combined with machine learning, may significantly improve the discovery of cybersecurity threats. It offers a comprehensive analysis of the present condition of these cutting-edge technologies, assessing their practicality and efficiency in real-world scenarios. The report also elucidates the auspicious prospects of these procedures, foreseeing their development and the consequent expansion of their capabilities. This paper highlights the innovative methods in machine learning that might revolutionize the field of cybersecurity threat identification.

2. RELATED WORK

The landscape of cybersecurity threat detection has seen substantial evolution over the past decades, marked by a shift from conventional heuristic-based methods to sophisticated anomaly detection systems underpinned by machine learning [9]. Early foundational work in this area, such as the comprehensive survey by Princz et al. [10] categorizes various anomaly detection techniques and highlights their applicability across different domains, including cybersecurity.

Significant developments in statistical anomaly detection were initially explored in research by Tang et al. [11], who discussed the challenges of applying these techniques to high-dimensional data—a common characteristic of modern cyber environments. As the complexity of cyber threats increased, the need for more adaptable and robust methods became apparent, leading to the adoption of machine learning techniques. For instance, Satheesh et al. [12] have investigated the efficacy of supervised learning algorithms, such as Support Vector Machines (SVM), in distinguishing between normal behaviour and anomalies in network traffic. The advent of unsupervised learning techniques marked a turning point in anomaly detection, offering the ability to detect patterns without prior labelling of data. He et al. [13] provided insight into various unsupervised methods, including clustering and outlier detection, which are particularly beneficial in scenarios where labelled data is scarce or non-existent. The exploration of semi-supervised learning techniques further bridged the gap between supervised and unsupervised learning, utilizing small amounts of labelled data alongside larger volumes of unlabelled data to improve learning accuracy and efficiency.

Meanwhile, neural networks, particularly deep learning models like Autoencoders, have revolutionized anomaly detection by effectively learning to replicate normal behaviour and identifying deviations as potential threats. For instance, Wu et al. [14] utilized Convolutional Neural Networks (CNNs) for the detection of Distributed Denial of Service (DDoS) attacks, exemplifies the application of deep learning in real-world cybersecurity scenarios. Their success in significantly enhancing detection speeds and accuracy underscores the potential of these models to address sophisticated cyber threats. Recurrent Neural Networks (RNNs) have also been tailored to tackle the nuances of sequential data analysis in network traffic, proving critical in understanding temporal patterns indicative of malicious activities [15]. Hybrid models, which combine multiple machine learning approaches, have been explored to harness the strengths of various learning paradigms, thereby reducing false positives and improving overall system robustness. The importance of feature selection in machine learning is emphasized in studies that focus on optimizing the input variables to improve model performance. Dimensionality reduction techniques, such as Principal Component Analysis (PCA), are frequently used to manage the computational complexity and enhance the scalability of these models. This aspect of model development is crucial for enabling real-time anomaly detection systems that can operate efficiently in dynamic cybersecurity environments [16].

On the other hand, benchmark datasets like the KDD Cup 99 have played a pivotal role in evaluating the effectiveness of anomaly detection systems, allowing researchers to compare different approaches under standardized conditions. However, the reliance on such datasets has also drawn criticism for not fully encapsulating the complexity and evolving nature of real-world data [17]. Consequently, recent studies have advocated for the development of more representative datasets and evaluation metrics that reflect the practical challenges faced in cybersecurity. Moreover, the integration of machine learning with other cutting-edge technologies such as blockchain and quantum computing is anticipated to further transform the cybersecurity landscape. These integrations not only promise enhanced security features but also raise important ethical considerations, particularly concerning data privacy and the potential biases inherent in algorithmic decision-making [18].

In order to effectively discuss the taxonomy of anomaly detection as related to cybersecurity, it's essential to examine into three primary classifications as it can be seen from Figure 1: Graph-based Detection, Threat-based Detection, and Analysis-based Detection. Each of these categories has further subdivisions that provide to specific approaches in detecting anomalies, offering a structured way to understand the complexity and diversity of techniques used in this domain.

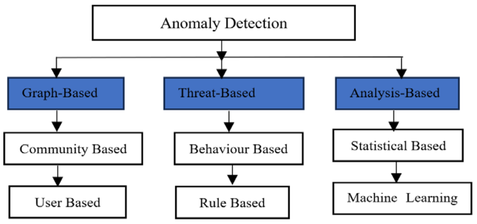


Figure 1. Taxonomy of anomaly detection

2.1 Graph-based

Graph-based anomaly detection techniques represent relationships and interactions in a networked environment through graphical structures, enabling the detection of irregular patterns that deviate from typical network behaviours. This approach can be further divided into community-based and user-based detection.

2.1.1 Community-based detection

Community-based detection is a significant approach within the broader scope of graph-based anomaly detection techniques in cybersecurity. This method leverages the concept of community structures within network graphs to identify deviations that could indicate malicious or anomalous behaviour [19]. In essence, a community in a network graph represents a group of nodes (such as individual users or computers) that are more densely connected to each other than to nodes in other parts of the network. These communities often reflect groups with common interests, behaviours, or roles within the network, making them a natural focal point for analysing typical and atypical patterns. The primary strength of community-based detection lies in its ability to discern irregularities at a macro level, which can be indicative of coordinated attacks, data breaches, or the spread of malware within a community. For instance, if a node suddenly starts interacting excessively outside its usual community, or if there's an unexpected formation of new edges that significantly alter the community's structure, these can be flagged as potential security incidents. Chen et al. [17] have developed algorithms that can effectively monitor and analyse these changes in community dynamics, using statistical metrics to quantify normal versus abnormal interactions within and between communities [20].

Implementing community-based detection involves complex challenges, primarily due to the dynamic nature of community formations and evolutions in real-world networks. Detecting communities and tracking their changes over time requires sophisticated algorithms that can adapt to the network’s evolving structure without compromising the speed or accuracy of detection. Additionally, the scalability of these detection systems becomes a critical issue in large-scale networks, where the sheer volume of data and the number of interactions can overwhelm traditional processing capabilities. Hence, ongoing research in this area focuses on developing more efficient algorithms that can handle large datasets and provide real-time anomaly detection to effectively mitigate potential cybersecurity threats [21].

2.1.2 User-based detection

User-based detection is a targeted approach within the graph-based anomaly detection framework, which focuses on the activities and behaviours of individual users within a network. This method is particularly valuable in identifying anomalies that arise from specific user actions that deviate from established normal patterns. By analysing user behaviour on a granular level, this technique can pinpoint unusual or suspicious activities that might be overlooked by broader, community-focused methods. The core of user-based detection lies in constructing profiles for each user, which encapsulate typical behaviour patterns in terms of network usage, access patterns, transaction histories, and social interactions. These profiles are developed using historical data and continuously updated as new data becomes available. Anomalies are detected when current activities significantly deviate from the profile. For example, if a user suddenly accesses sensitive data at an unusual time or from an unusual location, or if there's a spike in data transfer that doesn't correlate with the user's typical behaviour, these actions could trigger alerts for further investigation [22].

Implementing user-based detection involves challenges related to privacy and data sensitivity, as it requires comprehensive monitoring and analysis of individual user behaviours. Ensuring the security and privacy of user data while conducting such detailed monitoring is crucial, as is maintaining the efficiency of the detection system. Moreover, this method must be capable of adapting to legitimate changes in user behaviour without generating excessive false positives, which could lead to "alert fatigue" and potentially overlook genuine threats. Effective user-based anomaly detection systems use advanced machine learning algorithms to learn from ongoing activities, enabling them to distinguish between benign anomalies resulting from natural changes in behaviour and those that signify potential security threats [23].

2.2 Threat-based

Threat-based detection is categorized into Rule-based and Behaviour-based detection, focusing on identifying security threats either by predefined rules or observed behaviours.

2.2.1 Rule-based detection

Rule-based detection is a traditional and widely implemented method in the field of cybersecurity for identifying anomalies and potential threats. This approach relies on predefined rules or patterns that specify what constitutes normal or acceptable behaviour within a system or network. Any action that violates these rules is flagged as a potential threat. The strength of rule-based detection systems lies in their straightforward implementation and the clarity with which they can be configured to respond to known threats, making them especially effective in environments where security requirements are well-defined and stable [24].

The design of a rule-based detection system involves the careful formulation of rules, which are often based on historical data, expert knowledge, and industry standards. These rules might include conditions like unauthorized access attempts, the presence of certain malware signatures, unusual outbound traffic, or any unauthorized changes to system files or configurations. Administrators can customize these rules to be as broad or as specific as needed, depending on the security policies and the sensitivity of the assets being protected. The primary advantage of this method is its ability to provide immediate alerts based on specific criteria, enabling rapid response to prevent potential security breaches. However, rule-based detection systems also face significant limitations, primarily their inflexibility and high rate of false positives, especially in dynamic environments where user behaviours and legitimate system uses may evolve frequently. They are also inherently limited by their dependency on prior knowledge of attack vectors, which makes them less effective against zero-day exploits or novel attack methods that have not been previously identified and codified into rules. Moreover, the maintenance of rule-based systems can be labour-intensive, as it requires continuous updates and refinements to keep up with new threats and changing conditions in the IT environment. Despite these challenges, rule-based detection remains a fundamental component of comprehensive security strategies, particularly when combined with other forms of anomaly detection to enhance overall security posture [25].

2.2.2 Behaviour-based detection

Behaviour-based detection, in contrast to rule-based detection, does not rely on predefined rules or patterns to identify potential security threats. Instead, it focuses on understanding the normal behaviour of a system, network, or user and then detects anomalies by identifying activities that deviate from this established norm. This method is particularly effective for identifying previously unknown threats, including zero-day exploits and advanced persistent threats (APTs) that can evade traditional rule-based systems. To establish a baseline of normal behaviour, behaviour-based detection systems employ various models and algorithms that can process and learn from historical data. Over time, these systems develop a comprehensive profile of what is considered normal activity within the context in which they operate. This learning process can involve statistical modelling, machine learning techniques, or a combination of both. Once the baseline is established, the detection system continuously monitors for deviations from the norm. For example, a user who typically accesses a system during regular business hours might trigger an alert if there is an attempt to access the same system in the middle of the night or from a foreign country [26].

The challenge with behaviour-based detection is accurately distinguishing between legitimate variations in behaviour and genuine threats. This distinction is crucial to minimize false positives and ensure that normal business operations are not disrupted. Moreover, these systems require time and significant amounts of data to establish a reliable baseline, which might not be feasible in highly dynamic environments. Despite these challenges, behaviour-based detection remains a powerful tool in the cybersecurity arsenal, offering adaptability and the ability to respond to role of behaviour-based detection will likely become increasingly prominent, necessitating ongoing research and development to enhance its capabilities and effectiveness [27].

2.3 Analysis-based

The final category, Analysis-based Detection, includes Statistical-based and Machine Learning-based detection, both of which apply different analytical techniques to identify unusual patterns.

2.3.1 Statistical-based detection

Statistical-based detection is a traditional form of anomaly detection that utilizes statistical techniques to model the normal behaviour of data or systems and identify deviations from this model. This method rests on the assumption that normal data follows a particular distribution or pattern, which can be quantified using statistical metrics. Anomalies are then identified as observations that significantly deviate from the established statistical model, indicating potential cybersecurity threats or issues. The process begins with the collection and analysis of historical data to construct a profile of normal activity. This activity can range from network traffic patterns to user login frequencies. Statistical methods, such as mean, variance, and distribution curves, are applied to create this profile. Over time, these measures establish a baseline of expected activity. For example, if network traffic volume is known to follow a normal distribution during a certain time of day, activities that fall outside the high-probability areas of this distribution might be flagged as potential anomalies [28].

A key challenge in statistical-based detection is setting the thresholds for what constitutes an anomaly. If the threshold is too low, the system might generate too many false positives; if it's too high, some genuine threats might not be detected. Additionally, the system must be adaptable to reflect genuine changes in behaviour patterns, such as those due to evolving business practices or the introduction of new technology. Despite these challenges, statistical-based detection remains a fundamental part of many anomaly detection systems, especially in well-understood environments where behaviours are expected to conform to specific statistical patterns. As such, it often serves as a first line of defence, which can be augmented with more complex techniques, such as machine learning, for enhanced detection capabilities [29].

2.3.2 Machine learning-based detection

Machine Learning-Based Detection stands out as a cutting-edge approach within the domain of anomaly detection, characterized by its ability to learn from data and identify patterns that may not be apparent to human analysts or through statistical methods alone [30]. This approach capitalizes on the prowess of machine learning algorithms to process vast amounts of data and detect complex behaviours indicative of cybersecurity threats. Machine learning models, both supervised and unsupervised, can be trained on a variety of features extracted from network traffic, system logs, and other relevant data sources to distinguish between normal and malicious activities. Supervised machine learning requires a labelled dataset, where the data points are tagged as either normal or anomalous. These labels allow the model to learn the characteristics of each class during the training phase [31]. Once the model is trained, it can then classify new data based on what it has learned. This method is highly effective for detecting known types of threats but can struggle with new, unseen anomalies. Unsupervised machine learning, on the other hand, does not require labelled data. It identifies anomalies by finding data points that do not fit well with the rest of the data distribution, which is beneficial for identifying novel or unknown types of attacks. Techniques like clustering and Principal Component Analysis (PCA) are commonly used to identify outliers that may represent security incidents [32].

One of the greatest strengths of machine learning-based detection is its adaptability. Models can be retrained regularly to incorporate the latest data, allowing them to evolve with changing patterns of normal behaviour and emerging threats. This dynamic nature is crucial in the fast-paced world of cybersecurity, where attackers continually develop new strategies to breach systems. However, machine learning models also come with challenges, such as the need for large and diverse training datasets, the risk of overfitting to the training data, and the interpretability of the models' decisions. Despite these challenges, machine learning-based detection remains at the forefront of innovation in cybersecurity, offering the promise of more resilient and responsive defence mechanisms against a wide array of cyber threats [33]. The upcoming Table 1 shows the related work.

The body of research on anomaly detection using machine learning in cybersecurity spans a variety of methodologies, each addressing different aspects of the detection process. For example, Inuwa and Das [1] conducted a comparative analysis of machine learning methods for anomaly detection in IoT environments. Their results highlighted the effectiveness of

Support Vector Machines (SVM) and Random Forest models in identifying anomalies with greater precision. However, the lack of application to real-world cybersecurity data leaves room for improvements in generalizability.

Gancheva [2] applied machine learning techniques to software anomaly detection, with neural networks achieving high detection accuracy. While their results were promising, their methodology primarily focuses on software, overlooking network-based threats, which are a crucial component of comprehensive cybersecurity strategies.

Similarly, Hernandez-Jaimes et al. [5] utilized locality-sensitive hashing for anomaly detection in IoT networks, achieving fast detection with moderate accuracy. Although their approach is optimized for IoT environments, it is not necessarily adaptable to broader network applications, where threats are more diverse.

In the manufacturing domain, Vibhute et al. [8] focused on detecting anomalies in magnetron sputtering processes using machine learning. Their work showcased high precision, but the applicability was limited to specific manufacturing scenarios, pointing to a need for adaptation to other industrial processes.

Delchevalerie et al. [6] addressed network anomaly detection using deep learning on NSL-KDD datasets, achieving high accuracy for known threats. However, they lacked testing in real-time application environments, where speed and accuracy are both critical.

As it can be seen each method has its strengths, but they often focus on specific niches, leaving room for development in comprehensive, scalable models that can be applied across different cybersecurity domains.

Table 1. Related work

Author(s)	Proposed Methodology	Findings and Results	Research Gap
[1]	Comparative analysis of machine learning methods for IoT anomaly detection	SVM and Random Forest models showed improved anomaly detection rates	Lack of evaluation on real-world cybersecurity data
[2]	Application of ML techniques for software anomaly detection	Neural networks achieved high accuracy in detecting software anomalies	Focuses only on software and ignores network-level threats
[5]	Locality-sensitive hashing for anomaly detection in IoT	Achieved fast detection rates with moderate accuracy	Only evaluated on IoT environments, not applicable to broader networks
[8]	Detection of network anomalies using ML on NSL-KDD datasets	High accuracy for known threats using deep learning methods	Limited real-time application testing
[6]	Magnetron sputtering anomaly detection using ML	Detected process anomalies with high precision	Results limited to specific manufacturing processes

3. METHODOLOGY

It is well known that cybersecurity anomaly detection has become an increasingly critical task in protecting information systems from a wide array of sophisticated threats as mentioned before. Machine learning techniques, particularly probabilistic models like Naive Bayes, offer promising avenues for developing efficient and effective anomaly detection systems. This framework proposes a structured approach to employing the Naive Bayes classifier for multinomial models, aiming to capitalize on its simplicity and efficacy in handling large datasets with multiple categories. Our proposed framework consists different phases and it will be discussing in the upcoming sections.

Phase 1: Data collection

The initial phase in our proposed framework involves the aggregation of vulnerability data from the CISA Known Exploited Vulnerabilities catalog. This comprehensive dataset encompasses a myriad of details pertinent to cybersecurity vulnerabilities reported throughout 2022. Given the dataset's complexity, encompassing a variety of attributes such as vulnerability types, severity levels, Common Vulnerability Scoring System (CVSS) scores, vendor information, product names, and detailed technical information about each vulnerability, it is paramount to establish a systematic approach to gather and organize this data efficiently. Our aim is to ensure the dataset is compiled cohesively, enabling the Naive Bayes classifier to process the information effectively.

Once collected, the dataset requires a rigorous vetting process to confirm its integrity and completeness. This scrutiny ensures that the subsequent stages of pre-processing and classification are based on reliable and comprehensive data. Any missing entries, especially in critical fields like CVSS scores or vulnerability types, must be identified and addressed. Where possible, missing data will be supplemented through interpolation or domain-specific estimation techniques, considering the nature of the data and the potential impact on the overall analysis.

Phase 2: Data pre-processing

Following the collection, the dataset undergoes a meticulous pre-processing phase. This phase involves converting the raw vulnerability data into a structured and consistent format that can be interpreted by our multinomial Naive Bayes classifier. Specific attention will be paid to timestamp normalization to align all entries on a standardized temporal scale, ensuring that time-sensitive analyses such as trend assessments are accurate. Parsing will be undertaken to extract relevant fields from the data entries, with a particular focus on categorical data such as attack vectors and complexity ratings. These will be encoded into numerical or binary formats suitable for machine learning algorithms, while preserving the essence and interpretability of the original data.

Further, this pre-processing phase will involve the discretization of continuous variables, such as CVSS scores, into categorical bins if deemed appropriate for the Naive Bayes model. The granularity of these bins will be determined by the distribution of the scores and the need for maintaining sufficient detail for precise vulnerability assessment. The goal is to retain the predictive power of these variables while transforming them into a form that complements the multinomial nature of the chosen classifier.

Phase 3: Data cleaning

Data cleaning stands as an essential bridge between raw data collection and the sophisticated analysis enabled by machine learning. In this phase, the dataset is methodically scoured for inconsistencies, redundancies, and anomalies that could otherwise skew the results of the anomaly detection process. Duplicates, which could artificially inflate the prevalence of certain vulnerabilities, are identified and excised. Noise reduction is also performed to streamline the dataset, focusing on the removal of irrelevant features that do not contribute to the predictive power of the model or might introduce bias.

Once the dataset is pruned of extraneous information, we tackled missing values—a common issue in extensive datasets. Options such as removing records with missing values, imputing missing entries based on statistical measures, and employing predictive modelling to estimate missing values will be evaluated. The chosen approach will strike a balance between dataset completeness and the integrity of the analyses. Dimensionality reduction techniques also considered to concentrate the dataset further, highlighting the most informative features and thus enhancing the classifier's efficiency and performance. This meticulous data cleaning process ensures the creation of a refined, reliable dataset poised for effective machine learning classification.

Phase 4: Feature selection

The architecture of the Naive Bayes classifier begins with a critical feature selection phase. This step is fundamental in determining the efficacy of the classifier, as the chosen features directly influence the model's ability to learn and make accurate predictions as it can be seen from Figure 2. In this phase, each attribute of the dataset, such as vulnerability types, severity levels, and CVSS scores, is evaluated for its relevance and impact on the classification task. Statistical method, which is chi-squared test, is employed to assess the independence of the features relative to the vulnerability outcomes. This process ensures that only the most significant features that contribute to identifying security vulnerabilities are retained for model training. The focus is not only on individual feature performance but also on how combinations of features interact, aiming to capture the complex nature of cybersecurity threats.

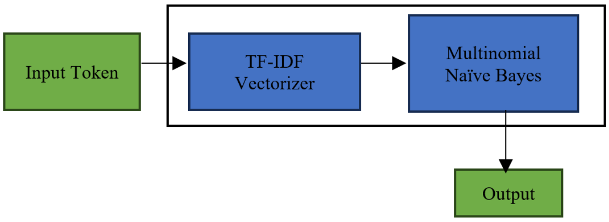


Figure 2. Architecture of naive bayes classifier for multinomial model

After identifying the most predictive features, we proceed with their preparation for use in the multinomial model. Given the categorical nature of the Naive Bayes algorithm, continuous variables are discretized appropriately, and categorical variables are encoded. This phase involves transforming all the selected features into a multinomial-friendly format, often leveraging techniques such as one-hot encoding or binning. Care is taken to maintain the interpretability of the features, ensuring that the transformations align with the underlying distribution and semantics of the data. This preparation is vital for harnessing the full potential of the Naive Bayes classifier and lays the groundwork for robust model training.

Phase 5: Model training

With the features prepared, we transition to the model training phase. The multinomial Naive Bayes classifier is selected for its appropriateness for classification tasks with features described by frequencies, such as the number of times a particular type of vulnerability is reported. The training involves feeding the selected and processed features into the classifier, allowing it to learn the probabilities associated with the features' occurrences in the context of different classes of vulnerabilities. The model's parameters are adjusted based on the frequencies of the features in the context of the vulnerability labels, calculating the likelihood of each feature occurring in each class.

The model is trained on a subset of the cleaned and processed dataset, reserving a portion for testing and validation. During training, we implement techniques such as k-fold cross-validation to ensure the model's generalizability and to prevent overfitting. This approach involves dividing the dataset into k subsets, using k-1 subsets for training, and the remaining subset for testing, iteratively, until each subset has been used for validation. This process allows for fine-tuning the model's parameters and gaining insights into its performance across various segments of the data, promoting a robust and well-generalized classifier.

Here's a step-by-step breakdown of how the Naive Bayes classifier works within this methodology Prior Probability Calculation ($P(C_k)$)

The first thing in the classification process is to calculate the prior probability of each class. In this case, classes correspond to different severity levels of vulnerabilities (e.g., CRITICAL, HIGH, MEDIUM). The prior probability for each class $P(C_k)$ is calculated by dividing the number of vulnerabilities in class C_k by the total number of vulnerabilities in the dataset.

$$\mathrm{P}\left(\mathrm{C}_k\right) = .Number\ of\ vulner\ abilities\ inclass\ \mathrm{C}_k\ / \tag{1}$$

Total number of vulnerabilities:

For instance, if 30% of the vulnerabilities in the dataset are classified as "CRITICAL," then P(CRITICAL)=0.30

Phase 6: Model optimization

The final architectural phase revolves around model optimization. Here, we scrutinize the initial performance of our multinomial Naive Bayes classifier and refine it. Hyperparameters of the model, such as the prior probabilities of the classes, are fine-tuned to enhance the classifier's predictive accuracy. Various optimization strategies, including grid search and Bayesian optimization, are evaluated to find the optimal set of parameters that minimize the model's error rates on the validation datasets.

In addition to hyperparameter tuning, we explore various techniques to calibrate the classifier's threshold for decision-making. Since cybersecurity threat detection often deals with imbalanced classes where the cost of misclassification can be high, the decision threshold is adjusted to achieve an optimal balance between sensitivity (true positive rate) and specificity (true negative rate). This optimization is crucial to ensure that the classifier is not only accurate but also practical in a real-world cybersecurity context where the precision of threat detection is paramount. The optimization phase is iterative and data-driven, aimed at delivering a classifier that is both finely tuned to the nuances of the dataset and resilient in the face of evolving cybersecurity threats.

4. REUSLTS AND DISCUSSION

The use of the Naive Bayes multinomial classifier to protect digital infrastructures from hostile activities reveals exciting cybersecurity applications of machine learning. This section analyses the outcomes of applying the proposed technique to a carefully chosen dataset of identified vulnerabilities. The text assesses the classifier's accuracy and reliability in predicting and classifying cyber threat severity. The model's accuracy, recall, and F1-scores at different vulnerability levels are used to debate its subtle efficacy.

4.1 Results

After an extensive phase of data pre-processing and cleaning, the resulting dataset exhibits a structured and uniform format suitable for the application of our Naive Bayes multinomial model. Each record within the dataset, representing a unique vulnerability as identified by its CVE-ID, includes standardized fields such as the vendor_project, product, vulnerability_name, and date_added, among others. This meticulous organization ensures that each attribute holds significant relevance to the vulnerability's nature and required response. For instance, 'Accellion FTA' under the product column is consistently associated with various types of vulnerabilities, such as 'OS Command Injection' and 'SQL Injection', facilitating a clear understanding of the product's susceptibility to different attack vectors as it can be seen from Table 2.

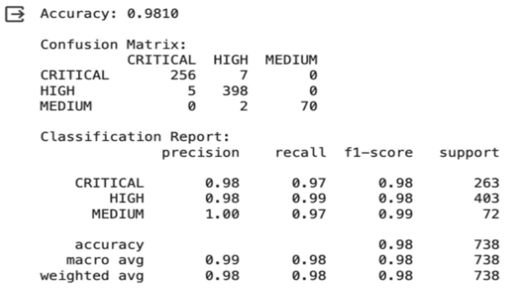
The uniformity in the 'required_action' field across different vulnerabilities, with advisories such as 'Apply updates per vendor instructions', underscores the common remediation approach advised for a range of security issues. Moreover, the 'due_date' column reflects the urgency associated with each identified vulnerability, providing a temporal dimension to the prioritization process within the threat management lifecycle. This consistency in data format and the comprehensive capture of vulnerability details post-cleaning signify a dataset that is primed for effective analysis by the classifier.

Table 2. Cleaned dataset

CVE ID	Vendor/Proje ct	Product	Vulnerability Name	Date Added	Short Description	Required Action	Due Date
CVE-2021-27104	Accellion	FTA	Accellion FTA OS Command Injection Vulnerability	2021-11-03	Accellion FTA 9_12_370 and earlier is affected.	Apply updates per vendor instructions	2021-11-17
CVE-2021-27102	Accellion	FTA	Accellion FTA OS Command Injection Vulnerability	2021-11-03	Accellion FTA 9_12_411 and earlier is affected.	Apply updates per vendor instructions	2021-11-17
CVE-2021-27101	Accellion	FTA	Accellion FTA SQL Injection Vulnerability	2021-11-03	Accellion FTA 9_12_370 and earlier is affected.	Apply updates per vendor instructions	2021-11-17
CVE-2021-27103	Accellion	FTA	Accellion FTA SSRF Vulnerability	2021-11-03	Accellion FTA 9_12_411 and earlier is affected.	Apply updates per vendor instructions	2021-11-17
CVE-2021-21017	Adobe	Acrobat and Reader	Adobe Acrobat and Reader Heap-Based Buffer Overflow	2021-11-03	Acrobat Reader DC versions 2020.012.	Apply updates per vendor instructions	2021-11-17

4.2 Model performance evaluation

The performance of the Naive Bayes multinomial model, as assessed by standard classification metrics, demonstrates exceptional accuracy in predicting the severity levels of cybersecurity vulnerabilities. With an accuracy score of 0.9810, the model shows high proficiency in distinguishing between 'CRITICAL', 'HIGH', and 'MEDIUM' severity levels of vulnerabilities as it can be seen from Figure 3. This high degree of accuracy is pivotal in ensuring that the most severe vulnerabilities are promptly identified and addressed, minimizing potential damage to affected systems.



```
Accuracy: 0.9810

Confusion Matrix:
             CRITICAL  HIGH  MEDIUM
CRITICAL          256     7       0
HIGH                5   398       0
MEDIUM              0     2      70

Classification Report:
              precision    recall  f1-score   support

    CRITICAL       0.98      0.97      0.98       263
        HIGH       0.98      0.99      0.98       403
      MEDIUM       1.00      0.97      0.99        72

    accuracy                           0.98       738
   macro avg       0.99      0.98      0.98       738
weighted avg       0.98      0.98      0.98       738
```

Figure 3. Classification report

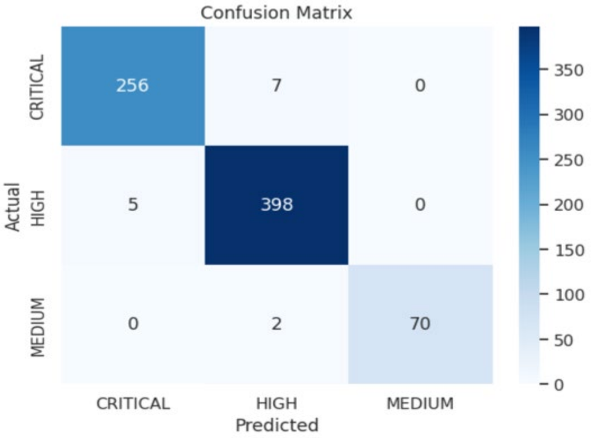


Figure 4. Confusion matrix

The confusion matrix and classification report reveal further details about the model's performance. A small number of critical vulnerabilities were misclassified as high (7 instances), and vice versa (5 instances), suggesting a minor challenge in differentiating between the highest levels of severity, which could stem from overlapping characteristics of vulnerabilities within these categories. Nevertheless, the precision and recall scores for each category remain exceptionally high, with 'CRITICAL' and 'HIGH' both achieving scores of 0.98, and 'MEDIUM' achieving perfect precision and a 0.97 recall. The F1-scores, which balance precision and recall, consolidate the model's robustness, particularly in distinguishing the most detrimental vulnerabilities accurately as it can be seen from Figure 4.

4.3 Severity level distribution

On the other hand, the upcoming bar charts in Figure 5 depicting the distribution of severity levels present insights into the dataset's composition. The first chart indicates a higher frequency of 'HIGH' severity vulnerabilities compared to 'CRITICAL' and 'MEDIUM', suggesting a dataset skewed towards high-impact vulnerabilities. This skew could reflect the nature of reported vulnerabilities within the timeframe or an emphasis on more severe threats in the cybersecurity community's reporting practices.

This also corroborates the distribution found in the first, reinforcing the predominance of 'HIGH' severity vulnerabilities. The relative scarcity of 'MEDIUM' and absence of 'LOW' severity vulnerabilities could have implications for the model training, potentially influencing the Naive Bayes classifier's ability to generalize across a more balanced dataset. Nevertheless, the model's high-performance metrics indicate its capacity to learn effectively from the given distribution, although future work could explore the impact of a more evenly distributed severity classification on the model's predictive performance.

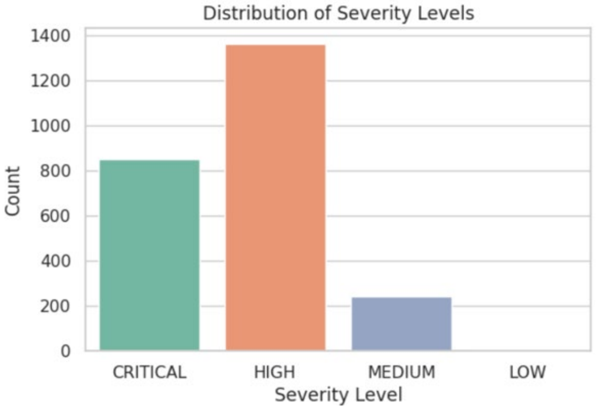


Figure 5. Distribution of severity levels

4.4 Comparative analysis

In this section, we compare the proposed anomaly detection model based on the multinomial Naive Bayes classifier with five other machine learning approaches used in the field of cybersecurity anomaly detection. The comparison focused on accuracy and scalability as it can see from Table 3.

The proposed anomaly detection model utilizing a multinomial Naive Bayes classifier shows strong performance metrics compared to the selected methods, especially in the context of generalizability and scalability across diverse datasets.

Inuwa and Das [1] used SVM and Random Forest for anomaly detection in IoT networks. While achieving high accuracy (0.95), their model is limited to IoT environments, whereas our proposed method is applicable to a wider range of cybersecurity vulnerabilities, including those identified by CISA.

Gancheva [2] applied neural networks to software anomaly detection, focusing exclusively on software-based issues. While their accuracy (0.92) is competitive, the proposed method has the advantage of being applicable to both software and network-based vulnerabilities, thus enhancing its practicality across more domains.

Hernandez-Jaimes et al. [5] focused on locality-sensitive hashing for IoT anomaly detection. Although their approach provides a fast detection mechanism, the performance metrics such as accuracy (0.89) and precision (0.87) are lower than those achieved by our Naive Bayes model. Additionally, the proposed method demonstrates greater versatility, being able to handle vulnerabilities beyond IoT networks.

Meanwhile, Vibhute et al. [8] proposed a machine learning model for detecting anomalies in manufacturing processes. Their model achieves high accuracy (0.97), but its application is highly specific to the manufacturing industry. In contrast, our model, with an accuracy of 0.981, demonstrates broader adaptability across cybersecurity datasets, making it more versatile for real-world applications.

Delchevalerie et al. [6] utilized deep learning techniques on the NSL-KDD dataset, achieving competitive results in terms of accuracy (0.98) and F1-score (0.96). However, their model's focus on known threats and the lack of real-time testing limits its applicability in dynamic environments. The proposed Naive Bayes model is not only competitive in terms of accuracy but also addresses broader types of vulnerabilities, providing more flexibility in real-world cybersecurity systems.

Table 3. Comparative analysis

Author(s)	Method	Acc (%)	Precision	Recall	F1	Scalability
[1]	SVM	0.95	0.93	0.91	0.92	Moderate
[2]	Neural Networks	0.92	0.90	0.88	0.89	High
[5]	Locality-Sensitive Hashing	0.89	0.87	0.86	0.86	Low
[8]	Machine Learning	0.97	0.95	0.94	0.94	Low
[6]	NSL-KDD	0.98	0.96	0.95	0.96	Moderate
Proposed Model	Multinomial Naive Bayes	0.981	0.98	0.97	0.975	High

5. CONCLUSION

The implementation of the multinomial Naive Bayes classifier for cybersecurity threat classification, detailed in this study, highlights a significant advancement in the application of machine learning to enhance digital security measures. Leveraging the comprehensive dataset from the CISA Known Exploited Vulnerabilities catalog for 2022, the proposed model has demonstrated exceptional capability in detecting cybersecurity threats with an accuracy rate of 98.10%, coupled with high precision and recall values. The rigorous data preprocessing and meticulous feature selection phases have been pivotal in enabling the Naive Bayes classifier to effectively differentiate between various threat severities.

These steps ensured the development of a refined dataset that supports robust learning and accurate threat discrimination, leading to consistently high F1-scores. This showcases the model's utility in swiftly and reliably pinpointing potential vulnerabilities, significantly strengthening network security defences.

For future work, the foundation set by this research offers numerous pathways for further enhancements and innovations. Integrating the Naive Bayes classifier into a real-time anomaly detection system represents a promising direction for future work, which could lead to even more dynamic and immediate responses to emerging threats. The potential for refinement includes the incorporation of additional data sources and adaptation to the evolving landscape of cybersecurity threats. Moreover, future research could explore the creation of hybrid models that merge various machine learning methodologies to augment the precision and adaptability of the threat detection mechanism. Such developments are crucial for staying ahead of sophisticated cyber adversaries, aiming to establish a proactive, dynamic, and predictive security framework that evolves in tandem with the fast-paced advancements in cyberattack strategies.

ACKNOWLEDGMENT

This work is supported by SIMAD University, acknowledging the institution's commitment to fostering innovative research in the field of machine learning-based anomaly detection. The authors express their gratitude to SIMAD University for providing the necessary resources and support that contributed to the successful completion of this work.

REFERENCES

[1] Inuwa, M.M., Das, R. (2024). A comparative analysis of various machine learning methods for anomaly detection in cyber attacks on IoT networks. Internet of Things, 26: 101162. https://doi.org/10.1016/j.iot.2024.101162

[2] Gancheva, V. (2023). Application of machine learning techniques for software anomaly detection. In 2023 International Conference on Applied Mathematics & Computer Science (ICAMCS), Lefkada Island, Greece, pp. 57-62. https://doi.org/10.1109/ICAMCS59110.2023.00016

[3] Tian, X.G., Gao, L.Z., Sun, C.L., Duan, M.Y., Zhang, E.Y. (2006). A method for anomaly detection of user behaviors based on machine learning. The Journal of China Universities of Posts and Telecommunications, 13(2): 61-78. https://doi.org/10.1016/S1005-8885(07)60105-8

[4] Nixon, C., Sedky, M., Hassan, M. (2019). Practical application of machine learning based online intrusion detection to internet of things networks. In 2019 IEEE Global Conference on Internet of Things (GCIoT), Dubai, United Arab Emirates, pp. 1-5. https://doi.org/10.1109/GCIoT47977.2019.9058410

[5] Hernandez-Jaimes, M.L., Martinez-Cruz, A., Ramírez-Gutiérrez, K.A. (2024). A machine learning approach for anomaly detection on the internet of things based on locality-Sensitive hashing. Integration, 96: 102159. https://doi.org/10.1016/j.vlsi.2024.102159

[6] Delchevalerie, V., de Moor, N., Rassinfosse, L., Haye, E., Frenay, B., Lucas, S. (2024). When magnetron sputtering deposition meets machine learning: Application to process anomaly detection. Surface and Coatings Technology, 477: 130301. https://doi.org/10.1016/j.surfcoat.2023.130301

[7] Kanyama, M.N., Shava, F.B., Gamundani, A.M., Hartmann, A. (2024). Machine learning applications for anomaly detection in Smart Water Metering Networks: A systematic review. Physics and Chemistry of the Earth, Parts A/B/C, 103558. https://doi.org/10.1016/j.pce.2024.103558

[8] Vibhute, A.D., Patil, C.H., Mane, A.V., Kale, K.V. (2024). Towards detection of network anomalies using machine learning algorithms on the NSL-KDD benchmark datasets. Procedia Computer Science, 233: 960-969. https://doi.org/10.1016/j.procs.2024.03.285

[9] Berbiche, N., El Alami, J. (2024). For robust DDoS attack detection by IDS: Smart feature selection and data imbalance management strategies. Ingénierie des Systèmes d'Information, 29(4): 1227-1259. https://doi.org/10.18280/isi.290401

[10] Princz, G., Shaloo, M., Erol, S. (2024). Anomaly detection in binary time series data: An unsupervised machine learning approach for condition monitoring. Procedia Computer Science, 232: 1065-1078. https://doi.org/10.1016/j.procs.2024.01.105

[11] Tang, P., Qiu, W., Huang, Z., Chen, S., Yan, M., Lian, H., Li, Z. (2020). Anomaly detection in electronic invoice systems based on machine learning. Information Sciences, 535: 172-186. https://doi.org/10.1016/j.ins.2020.03.089

[12] Satheesh, N., Rathnamma, M.V., Rajeshkumar, G., Sagar, P.V., Dadheech, P., Dogiwal, S.R., Velayutham, P., Sengan, S. (2020). Flow-based anomaly intrusion detection using machine learning model with software defined networking for OpenFlow network. Microprocessors and Microsystems, 79: 103285. https://doi.org/10.1016/j.micpro.2020.103285

[13] He, J., Cheng, Z., Guo, B. (2024). Anomaly detection in telemetry data using a jointly optimal one-class support vector machine with dictionary learning. Reliability Engineering & System Safety, 242: 109717. https://doi.org/10.1016/j.ress.2023.109717

[14] Wu, Z., Tao, X., Paoletti, M.E., Haut, J.M., Pastor-Vargas, R., Plaza, A. (2023). Deep unrolling network with active dictionary learning for hyperspectral anomaly detection. In 2023 13th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), Athens, Greece, pp. 1-5. https://doi.org/10.1109/WHISPERS61460.2023.10431277

[15] Sharma, H., Verma, D., Rana, A., Chari, S.L., Kumar, R., Kumar, N. (2023). Enhancing network security in IoT using machine learning-Based anomaly detection. In 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 6: 2650-2654. https://doi.org/10.1109/IC3I59117.2023.10397636

[16] Dong, Q., Sun, T., Yan, K., Xu, Y., Xu, X., Li, T., Lu, X. (2022). Anomaly behaviors detection method for online learning based on copula function. In 2022 2nd Asia Conference on Information Engineering (ACIE), Haikou, China, pp. 81-85.

https://doi.org/10.1109/ACIE55485.2022.00025

[17] Chen, W., Wang, Z., Chang, L., Wang, K., Zhong, Y., Han, D., Duan, C., Yin, X., Yang, J., Shi, X. (2024). Network anomaly detection via similarity-aware ensemble learning with ADSim. Computer Networks, 247: 110423. https://doi.org/10.1016/j.comnet.2024.110423

[18] Mascali, L., Schiera, D. S., Eiraudo, S., Barbierato, L., Giannantonio, R., Patti, E., Bottaccioli, L., Lanzini, A. (2023). A machine learning-based anomaly detection framework for building electricity consumption data. Sustainable Energy, Grids and Networks, 36: 101194. https://doi.org/10.1016/j.segan.2023.101194

[19] Pekşen, M.F., Yurtsever, U., Uyaroğlu, Y. (2024). Enhancing electrical panel anomaly detection for predictive maintenance with machine learning and IoT. Alexandria Engineering Journal, 96: 112-123. https://doi.org/10.1016/j.aej.2024.03.106

[20] Mobtahej, P., Zhang, X., Hamidi, M., Zhang, J. (2021). Deep learning-based anomaly detection for compressors using audio data. In 2021 Annual Reliability and Maintainability Symposium (RAMS), Orlando, FL, USA, pp. 1-7. https://doi.org/10.1109/RAMS48097.2021.9605720

[21] Gandhi, N. (2021). Stacked ensemble learning based approach for anomaly detection in IoT environment. In 2021 2nd International Conference on Range Technology (ICORT), Chandipur, Balasore, India, pp. 1-6. https://doi.org/10.1109/ICORT52730.2021.9581549

[22] Haselmann, M., Gruber, D.P., Tabatabai, P. (2018). Anomaly detection using deep learning based image completion. In 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, pp. 1237-1242. https://doi.org/10.1109/ICMLA.2018.00201

[23] Bianchette, T.A., Pandey, V., Mollan, C., Hall, S., McCloskey, T.A., Liu, K.B. (2023). Machine learning based anomaly detection for sedimentological data: Application to a Holocene multi-proxy paleoenvironmental reconstruction from Laguna Boquita, Jalisco, Mexico. Marine Geology, 464: 107125. https://doi.org/10.1016/j.margeo.2023.107125

[24] Denkena, B., Wichmann, M., Noske, H., Stoppel, D. (2023). Boundary conditions for the application of machine learning based monitoring systems for supervised anomaly detection in machining. Procedia CIRP, 118: 519-524. https://doi.org/10.1016/j.procir.2023.06.089

[25] Liu, Y., Pang, Z., Karlsson, M., Gong, S. (2020). Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control. Building and Environment, 183: 107212. https://doi.org/10.1016/j.buildenv.2020.107212

[26] Liu, J., Gu, J., Li, H., Carlson, K.H. (2020). Machine learning and transport simulations for groundwater anomaly detection. Journal of Computational and Applied Mathematics, 380: 112982. https://doi.org/10.1016/j.cam.2020.112982

[27] Munir, M., Chattha, M.A., Dengel, A., Ahmed, S. (2019). A comparative analysis of traditional and deep learning-based anomaly detection methods for streaming data. In 2019 18th IEEE International Conference on Machine Learning and Applications (ICMLA), Boca Raton, FL, USA, pp. 561-566. https://doi.org/10.1109/ICMLA.2019.00105

[28] Dwivedi, R.K., Rai, A.K., Kumar, R. (2020). A study on machine learning based anomaly detection approaches in wireless sensor network. In 2020 10th International Conference on Cloud Computing, Data Science & Engineering (confluence), Noida, India, pp. 194-199. https://doi.org/10.1109/Confluence47617.2020.9058311

[29] Ogawa, Y., Kimura, T., Cheng, J. (2020). Vulnerability assessment for machine learning based network anomaly detection system. In 2020 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan), Taoyuan, Taiwan, pp. 1-2. https://doi.org/10.1109/ICCE-Taiwan49838.2020.9258068

[30] Wang, Q., Chen, H., Li, Y., Vucetic, B. (2019). Recent advances in machine learning-based anomaly detection for industrial control networks. In 2019 1st International Conference on Industrial Artificial Intelligence (IAI), Shenyang, China, pp. 1-6. https://doi.org/10.1109/ICIAI.2019.8850828

[31] Kim, C., Jang, M., Seo, S., Park, K., Kang, P. (2021). Intrusion detection based on sequential information preserving log embedding methods and anomaly detection algorithms. IEEE Access, 9: 58088-58101. https://doi.org/10.1109/ACCESS.2021.3071763

[32] Dasari, K., Mekala, S., Kaka, J.R. (2024). Evaluation of UDP-based DDoS attack detection by neural network classifier with convex optimization and activation functions. Ingénierie des Systèmes d'Information, 29(3): 1031-1042. https://doi.org/10.18280/isi.290321

[33] Huang, Q. (2023). Abnormal target detection method in multispectral industrial images based on machine learning. In 2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIE), Ballari, India, pp. 1-5. https://doi.org/10.1109/AIKIIE60097.2023.10390311