Al-Rafidain Journal of Engineering Sciences

Journal homepage https://rjes.iq/index.php/rjes

ISSN 3005-3153 (Online)

An Advanced Security System for Detecting Cyber-attacks in Computer Networks Based on Machine Learning

Majida Hamid Hamzah

University of Samarra

ARTICLE INFO

Article history:
Received 12 January 2026
Revised 12 January 2026
Accepted 16 January 2026
Available online 17 January 2026

Keywords:
Cybersecurity,
Intrusion Detection System,
Machine Learning,
Cyberattack Detection,
Network Security

ABSTRACT

Computer networks have become a fundamental component of modern digital infrastructures across various sectors, including government institutions, banking systems, telecommunications, and electronic commerce. With the rapid expansion of network usage, cyberattacks have increased significantly in both frequency and complexity, posing serious challenges to traditional security systems. Conventional intrusion detection systems, which rely mainly on predefined rules and signature-based mechanisms, have demonstrated clear limitations in detecting advanced and previously unknown attacks, in addition to high false alarm rates. This research presents the design and implementation of an advanced security system for detecting cyberattacks in computer networks based on machine learning techniques. The proposed system analyzes network traffic data and classifies activities into normal and malicious behaviors using intelligent learning models. Several machine learning algorithms, including Support Vector Machine (SVM), Random Forest, and Artificial Neural Networks (ANN), were applied and evaluated to assess their effectiveness in cyberattack detection. The system was developed following a structured methodology that includes data collection, preprocessing, feature extraction, model training, and performance evaluation. Standard evaluation metrics such as accuracy, precision, recall, and F1-score were used to measure system performance. Experimental results demonstrated that machine learning–based approaches significantly improve detection accuracy and reduce false alarm rates compared to traditional methods. Among the evaluated algorithms, the Random Forest model achieved the best overall performance, particularly in terms of recall and F1-score.

The findings of this study confirm the effectiveness of integrating machine learning techniques into intrusion detection systems and highlight their potential in enhancing cybersecurity capabilities. The proposed system contributes to improving early attack detection, increasing system reliability, and addressing emerging cybersecurity challenges in modern network environments.

Corresponding author E-mail address: Majida.h.hemza@uosamarra.ed.iq
https://doi.org/10.61268/qgwn1w12

1. Introduction

Computer networks have witnessed rapid development in recent decades and have become the backbone of digital infrastructures across various sectors, including governmental institutions, the banking sector, telecommunications, and electronic commerce. With the extensive expansion in the use of networks, the frequency and complexity of cyberattacks have increased significantly, making network security one of the most critical challenges facing modern organizations. Cyberattacks are no longer limited to predictable traditional methods; rather, they have become more intelligent, adaptive, and capable of bypassing conventional security systems based on fixed signatures and predefined rules.In this context, traditional intrusion detection systems have shown clear limitations in dealing with new and unknown attacks, in addition to their high rates of false alarms, which reduce their practical effectiveness. As a result, there has been growing interest in machine learning techniques as a promising approach in the field of cybersecurity, due to their ability to analyze large volumes of network data, discover hidden patterns, and adapt to evolving threats.

1.1 Research Problem

Modern computer networks are facing a continuous escalation in cyberattacks in terms of both volume and complexity, alongside the widespread reliance on digital systems across various sectors. Despite the development of traditional security mechanisms, these systems suffer from clear limitations in detecting advanced and previously unknown attacks, in addition to high false alarm rates and weak adaptability to evolving threats. Therefore, the research problem lies in the need to develop an intelligent security system based on machine learning techniques that is capable of analyzing network traffic and detecting cyberattacks with higher accuracy and efficiency compared to traditional approaches.

1.2 Significance of the Research

1- Enhances the level of cybersecurity in computer networks.
2- Highlights the role of machine learning techniques in improving cyberattack detection accuracy.
3- Reduces false alarm rates, thereby increasing the effectiveness of security systems.
4- Provides an advanced security model applicable to various network environments.
5- Keeps pace with modern developments in the field of information security.

1.3 Research Objectives

1- To design an advanced security system for detecting cyberattacks in computer networks.
2-To analyze network traffic using machine learning algorithms.
3-To compare the performance of different machine learning algorithms in detecting cyberattacks.
4-To improve detection accuracy and reduce error rates and false positives.
5-To evaluate the efficiency of the proposed system using standard performance metrics.

1.4 Research Hypotheses

1-There is a statistically significant relationship between the use of machine learning techniques and improved accuracy in cyberattack detection.
2-A machine learning-based security system achieves better performance in detecting cyberattacks than traditional security systems. (Alpaydin, 2014)
3-The quality of network traffic data used for model training has a direct impact on the effectiveness of the security system.

1.5 Research Methodology

This research adopts a descriptive-analytical approach to study concepts related to network security and cyberattacks, in addition to an

experimental approach to design and develop the proposed security system. Standard network traffic datasets are used to train machine learning models, and the system's performance is evaluated using metrics such as accuracy, recall, precision, and error rate.

1.6 Research Limitations

1- The practical implementation is limited to benchmark or experimental network traffic datasets.
2- The study focuses on specific types of cyberattacks.
3- The research results depend on the selected machine learning algorithms and testing environment.
4- Legal and regulatory aspects of cybersecurity are beyond the scope of this research.

1.7 Previous studies

1- Alpaydin (2014) :This study aimed to present the theoretical foundations of machine learning techniques and their applications in analyzing complex data, including computer network data. The findings confirmed that classification algorithms possess a strong capability to uncover hidden patterns within large datasets, making them effective tools for distinguishing between normal and abnormal behavior. This study represents an important theoretical reference supporting the adoption of machine learning in intrusion detection systems, as it provides the conceptual framework upon which the current research builds its intelligent cyberattack detection approach.

2- Bishop (2006) This study focused on pattern recognition models based on machine learning, particularly Artificial Neural Networks (ANN) and Support Vector Machines (SVM), and their applications in information security. The results demonstrated that advanced statistical models can classify highly complex data with high accuracy compared to traditional techniques. The study highlighted the effectiveness of intelligent models in improving cyberattack detection accuracy, which aligns with the current research that employs SVM and ANN within the proposed security system.

3- Bocca, Mellia, and Meo (2019) This study investigated the development of an intelligent system for detecting network faults and cyberattacks based on machine learning techniques and network traffic analysis. The results showed that machine learning-based models are capable of early detection of anomalous activities with high accuracy while significantly reducing false alarm rates. The importance of this study lies in its use of real network traffic data, which enhances the reliability of its findings and provides a practical foundation supporting the current research in designing an effective and deployable security system.

2. Computer Networks and Information Security

Computer networks are considered one of the fundamental pillars of information technology infrastructure, as they enable the interconnection of devices and the efficient exchange of data and resources within and across organizations. The rapid development of networking technologies has significantly contributed to supporting digital transformation and expanding the scope of electronic services. However, this advancement has also led to an increase in the scale and sophistication of security threats. Information security refers to a set of technical and organizational measures and policies aimed at protecting data from unauthorized access, modification, or destruction, while ensuring its confidentiality, integrity, and availability. With the growing prevalence of cyberattacks, achieving network and information security has become a critical necessity to maintain system

continuity and safeguard digital assets. (Cisco Systems, 2020)

2.1 Concept of Computer Networks and Their Types

Computer networks refer to a group of interconnected devices and systems that are linked together through wired or wireless communication media for the purpose of exchanging data and information and sharing resources such as files, applications, printers, and databases. Networks are a fundamental component of modern information technology environments, as they provide fast communication, efficient resource management, and support for various electronic applications and services. Computer networks can be classified according to several criteria, most notably geographical scope, connection type, and purpose of use. Based on geographical scope, networks are divided into Local Area Networks (LAN), which operate within a limited area such as offices or buildings and are characterized by high speed and relatively low cost. Metropolitan Area Networks (MAN) cover a wider area such as a city and interconnect multiple local networks. Wide Area Networks (WAN) span large geographical distances and may extend across countries or continents, with the Internet being the most prominent example. In terms of connection type, networks are classified into wired networks, which rely on physical media such as optical fiber cables, and wireless networks, which depend on radio waves, such as Wi-Fi networks. Networks can also be categorized according to their purpose into private and public networks. This diversity enables networks to meet the varying needs of organizations and users, depending on performance and security requirements. (Géron, 2019)

2.2 Concept of Cybersecurity

Cybersecurity is defined as a set of policies, procedures, and technologies aimed at protecting computer systems, networks, data, and applications from digital threats and cyberattacks. This includes protecting information from unauthorized access, manipulation, or destruction, as well as ensuring service continuity and the integrity of digital infrastructures. Cybersecurity has become an essential element due to the increasing reliance on digital technologies and communication networks across various aspects of life. The concept of cybersecurity is based on fundamental principles known as the information security triad: Confidentiality, which ensures that information is accessible only to authorized individuals; Integrity, which maintains the accuracy and reliability of information by preventing unauthorized modification; and Availability, which ensures that information and systems can be accessed when needed. With the evolution of cyberattack techniques, the scope of cybersecurity has expanded to include risk management, early threat detection, incident response, and recovery. Cybersecurity is considered a shared responsibility that requires the integration of technical, organizational, and human aspects through the implementation of effective security policies, user awareness and training, and the use of advanced technologies such as intrusion detection systems and machine learning.

2.3 Types of Cyberattacks

Cyberattacks vary according to their objectives and methods of execution, and they are considered among the most serious threats facing computer networks and information systems. The most prominent types of cyberattacks include the following:

1-Denial of Service Attacks (DoS / DDoS): These attacks aim to disrupt the operation of systems or networks by overwhelming them with a massive number of requests, leading to resource exhaustion and preventing legitimate users from accessing services.

2-Malware Attacks: These include viruses, worms, Trojan horses, and ransomware,

which are designed to damage data, steal sensitive information, or gain unauthorized control over targeted systems. (Goodfellow et al., 2016)

3-Phishing Attacks: These attacks rely on deceiving users through fraudulent emails or fake websites in order to steal sensitive information such as passwords or banking details.

4-Hacking Attacks: These involve exploiting security vulnerabilities in systems or networks to gain unauthorized access to data or resources.

5-Man-in-the-Middle (MITM) Attacks: In these attacks, the attacker intercepts or manipulates communications between two parties without their knowledge, allowing the theft or alteration of transmitted data.

6-Social Engineering Attacks: These attacks exploit the human factor rather than technical vulnerabilities, using psychological manipulation to obtain confidential information from users. These types of cyberattacks highlight the need for advanced security systems that rely on modern techniques, such as machine learning, to enable early detection of abnormal activities and reduce cyber risks.

2.4 Intrusion Detection Systems (IDS)

Intrusion Detection Systems (IDS) are considered one of the fundamental pillars of network and information security, as they play a vital role in monitoring data traffic and detecting abnormal activities or malicious behaviors that may indicate cyberattacks. With the continuous increase in the complexity of attacks and the diversity of intrusion techniques, traditional security mechanisms alone have become insufficient, making IDS an essential complementary component for strengthening network security. DS operate by analyzing network traffic or system logs in order to distinguish between normal behavior and unauthorized or malicious activities, and then generating alerts when intrusion attempts or suspicious actions are detected. The importance of these systems lies in their ability to provide early detection of attacks, enabling organizations to respond quickly and minimize potential damage. In addition, IDS contribute to offering a comprehensive view of the network's security status and help identify vulnerabilities that attackers may exploit. With the advancement of digital technologies, intrusion detection systems have undergone significant development, evolving from signature-based and rule-based systems to more intelligent approaches that rely on behavioral analysis and machine learning techniques (Goodfellow et al., 2016).

2.5 Types of Intrusion Detection Systems (IDS)

Intrusion Detection Systems (IDS) can be classified based on their mode of operation, monitoring location, and analysis method. This diversity aims to meet different security requirements within network environments. The main types of IDS include the following:

1-Network-Based Intrusion Detection Systems (NIDS): These systems monitor overall network traffic by analyzing packets transmitted between devices. They are effective in detecting attacks that target the network as a whole, such as denial-of-service attacks and network scanning activities. (Han et al., 2011)

2-Host-Based Intrusion Detection Systems (HIDS): These systems operate on individual hosts or servers by monitoring system logs, files, and local activities. They are particularly effective in detecting insider attacks or unauthorized changes to system files.

3-Signature-Based Intrusion Detection Systems: These systems rely on comparing current activities against a database of known attack signatures. They provide high accuracy in detecting known attacks but have limited capability in identifying new or previously unknown threats. (Hassan et al., 2023)

4-Anomaly-Based Intrusion Detection Systems: These systems build a model of normal network or system behavior and detect deviations from this baseline. They are capable of identifying novel attacks; however, they may suffer from high false positive rates.

2.6 Challenges and Limitations of Intrusion Detection Systems (IDS)

Intrusion Detection Systems (IDS) face several challenges and limitations that can affect their efficiency and accuracy, particularly in light of the continuous evolution of cyberattacks and the increasing complexity of network environments. The most significant challenges include the following:

1- High False Positive Rates: False alarms are among the most common issues in IDS, as normal activities may be incorrectly classified as attacks, leading to confusion for security teams and increased operational workload.

2- Limited Ability to Detect New Attacks: Signature-based systems suffer from limitations in identifying unknown or advanced attacks, as they rely on predefined rules and signatures that may not cover emerging threats.

3- Complexity and Configuration Difficulty: IDS deployment and configuration require a high level of technical expertise. Improper configuration can result in reduced system effectiveness or increased error rates.

4- Performance and Resource Consumption: Real-time analysis of large volumes of network traffic may consume significant computational resources, potentially impacting network or system performance.

5- Scalability Issues: Some IDS solutions face difficulties in scaling to large or dynamic environments, especially with the rapid growth of data volume and multiple monitoring points. These challenges highlight the need for more intelligent intrusion detection systems that leverage advanced techniques such as machine learning and deep learning to overcome existing limitations and improve the accuracy and efficiency of cyberattack detection. (Liao et al., 2013)

2.7 Difference Between Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS)

Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS) are essential tools in the field of network security; however, each serves a different function and operates using a distinct approach. The differences between them can be summarized as follows:

1-An IDS focuses on monitoring network traffic or system activities to detect abnormal behavior or potential cyberattacks and then generates alerts to notify security administrators without directly interfering with data flow. (Mitchell, 1997)

In contrast, an IPS not only detects attacks but also actively prevents them in real time by blocking malicious traffic or terminating suspicious sessions.

2-Mode of Operationist is typically deployed as a passive system, as it only performs detection and alerting.

IPS, on the other hand, operates as an active system that takes immediate actions such as dropping packets or blocking suspicious IP addresses.

3-Deployment Location: IDS is usually deployed out-of-band, meaning it is not placed directly in the data path, which minimizes its impact on network performance.IPS is deployed inline within the data path, allowing it to prevent attacks but potentially affecting performance if not properly configured.

4-Risk and Impact's poses no direct risk to legitimate traffic but relies on human intervention to respond to detected threats. IPS may inadvertently block legitimate traffic if misconfigured, which could disrupt normal network operations.

2.8 Machine Learning in Cybersecurity

Machine learning is considered one of the most prominent modern technologies that has brought a fundamental transformation to the field of cybersecurity, due to its ability to analyze large volumes of data and identify complex patterns that are difficult for traditional methods to detect. With the continuous evolution of cyberattacks in terms of techniques, speed, and stealth, security solutions based on static rules and signatures have become insufficient to address advanced threats, which has necessitated the adoption of intelligent techniques capable of learning and adapting to dynamic environments. Machine learning in cybersecurity relies on training computational models using historical data from network traffic or system logs in order to distinguish between normal and malicious behavior. Various algorithms, such as classification, clustering, and anomaly detection algorithms, are employed in multiple applications, including intrusion detection systems, malware detection, abnormal behavior analysis, and early cyberattack detection. (Scikit-Learn Developers, 2020)

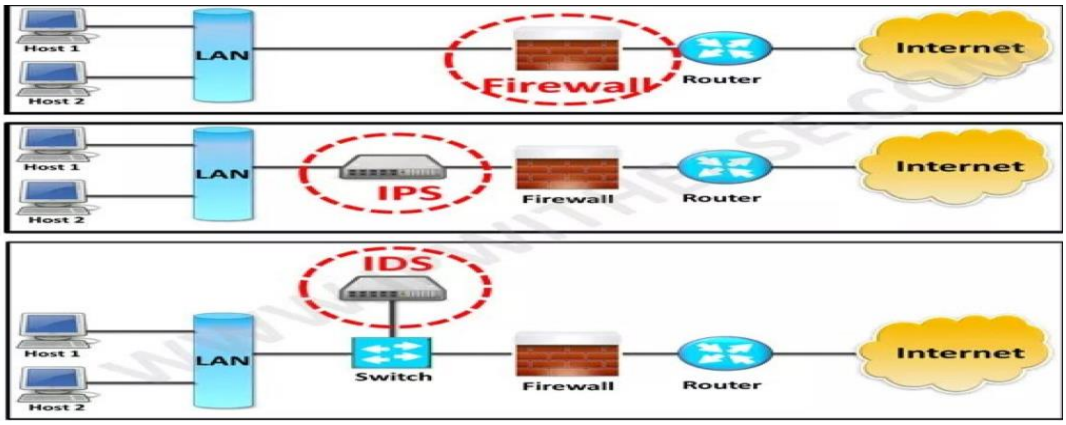

Figure 1

2.9 Classification Algorithms

Classification algorithms are among the most widely used machine learning techniques in the field of cybersecurity. They aim to assign data instances to predefined classes, such as distinguishing between normal and malicious behavior in computer networks. These algorithms rely on labeled training data containing known classes, enabling the model to learn the patterns associated with each class and apply them to classify new, unseen data.

Some of the most prominent classification algorithms used in cybersecurity include:

1-Support Vector Machine (SVM): This algorithm works by finding an optimal separating hyperplane between different classes in the feature space. It is known for its high accuracy in classifying complex data, particularly in high-dimensional spaces.

2-Decision Tree: Decision trees represent decisions in a hierarchical tree structure composed of nodes and branches, making them

easy to understand and interpret. They are commonly used to analyze network behavior.

3-Random Forest:This algorithm consists of a collection of decision trees and improves classification accuracy while reducing the risk of overfitting.

4-Artificial Neural Networks (ANN): These networks simulate the functioning of the human brain in processing information and are highly effective in detecting complex patterns within network data (Sommer & Paxson, 2010) .

2.10 Clustering Algorithms

Clustering algorithms are a type of unsupervised machine learning techniques that aim to group data into similar clusters based on shared characteristics, without relying on predefined class labels. These algorithms are particularly important in the field of cybersecurity due to their ability to discover unknown patterns and anomalous behaviors within network traffic.In cybersecurity applications, clustering algorithms are used to detect abnormal activities and identify new or previously unseen attacks by grouping similar behaviors and separating them from normal patterns. Some of the most widely used clustering algorithms include:

1-K-Means Algorithm: One of the simplest and most commonly used clustering algorithms, K-Means partitions data into a predefined number of clusters based on the distance between data points and cluster centroids. (Stallings, 2018)

2-Hierarchical Clustering: This approach builds a hierarchical structure of clusters, allowing the analysis of data relationships at multiple levels without the need to specify the number of clusters in advance.

3-DBSCAN Algorithm: This algorithm groups data points based on density and is well known for its ability to detect clusters with irregular shapes and to effectively handle noise and outliers. Clustering algorithms contribute to enhancing intrusion detection systems by enabling the discovery of unknown attacks and reducing reliance on fixed signatures, making them an effective tool for building advanced machine learning-based security systems.

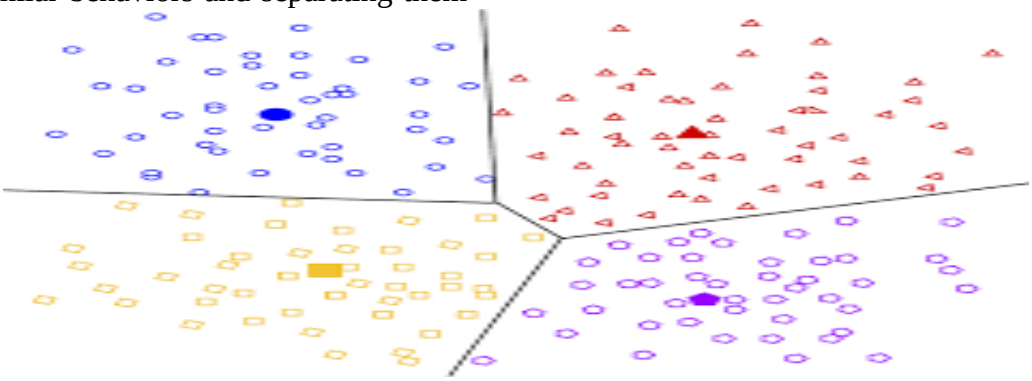

Figure 2

2.11 Applications of Machine Learning (ML) in Cyberattack Detection

Machine learning applications represent one of the most significant recent advances in the field of cyberattack detection, as they provide a high capability to analyze large volumes of data and identify abnormal patterns and behaviors that may indicate the presence of an attack. The use of machine learning techniques has helped overcome the limitations of traditional rule-based and signature-based approaches, particularly in addressing advanced and previously unknown attacks. (Stallings, 2018)

One of the most prominent applications of machine learning in attack detection is intelligent intrusion detection systems (IDS), where classification and clustering algorithms are used to analyze network traffic and

accurately distinguish between normal and malicious activities. In addition, anomaly detection techniques are employed to identify unusual behaviors that may represent zero-day attacks.

Other applications of machine learning include malware detection through analyzing file features and program behavior during execution, as well as phishing attack detection by examining email content and suspicious link patterns. Machine learning is also widely used in User and Entity Behavior Analytics (UEBA) to detect insider threats and attacks that exploit compromised user accounts.

3. Model Design

Model design represents the fundamental stage in building the proposed security system for cyberattack detection based on machine learning techniques. During this stage, the overall architecture of the system, its operational mechanism, and its main components are defined. This section aims to clarify the steps involved in designing the model in a way that ensures the highest possible level of accuracy and efficiency in detecting cyberattacks, while taking performance requirements and scalability into consideration.The proposed model relies on collecting network traffic data from reliable sources, whether real-world data or benchmark datasets commonly used in cybersecurity research. This is followed by a data preprocessing phase, which includes data cleaning, handling missing values, and extracting important features that contribute to improving the performance of machine learning algorithms. In the next stage, appropriate machine learning algorithms, such as classification or clustering algorithms, are selected and trained using the available training data. (Mitchell, 1997) Model hyperparameters are also tuned to achieve optimal performance. The model design is based on a flexible architecture that allows the integration of new algorithms or the updating of existing models to keep pace with evolving cyber threats.

3.1 System Architecture

The architecture of the proposed cyberattack detection system based on machine learning relies on a layered design that ensures clear organization of components, ease of integration, and scalability. The system architecture consists of several main layers that work together to achieve effective attack detection, as described below:

1-Data Collection Layer: This layer represents the starting point of the system, where network traffic data are collected from various sources, such as network devices, system logs, or benchmark datasets. These data provide the foundation for the analysis and detection processes. (Mitchell, 1997)

2-Preprocessing Layer: This layer is responsible for cleaning and preparing the data, including removing duplicate records, handling missing values, normalizing data, and converting them into a suitable format for input to machine learning models.

3-Feature Extraction Layer: In this layer, the most influential features are identified and extracted from network data, which contributes to improving model performance and reducing computational complexity.

4-Machine Learning Layer: This layer represents the core of the system, where selected machine learning algorithms are applied to classify or cluster data and detect malicious or abnormal activities. (Mitchell, 1997)

5- Detection and Response Layer: This layer analyzes the outputs of the models and generates alerts when potential attacks are detected, with the possibility of integrating with other systems to initiate appropriate response actions.

3.2 Data Sources

Data sources are considered a fundamental component in building an effective machine learning–based cyberattack detection system, as the efficiency and accuracy of the model largely depend on the quality of the data used for

training and testing. In this research, the data sources consist of network traffic data that reflect both normal and malicious behavior within different network environments. The data sources include benchmark datasets that are widely used in cybersecurity research, such as network traffic datasets containing multiple types of cyberattacks and normal activities. These datasets are characterized by the availability of accurately labeled data, which contributes to reliable training and evaluation of machine learning models.

In addition, real network traffic data can be utilized, collected from actual network environments using network monitoring tools, such as system logs or traffic analysis tools. These data provide a realistic representation of security threats and user behavior, thereby enhancing the model's ability to operate effectively in real-world scenarios. All collected data undergo preprocessing procedures, including data cleaning, feature extraction, and class balancing, in order to improve data quality and reduce bias

3.3 Preprocessing Steps

Data preprocessing is a fundamental stage in building machine learning–based cyberattack detection systems, as it directly contributes to improving model accuracy and efficiency. The main preprocessing steps include the following:

1-Data Cleaning:This step involves removing duplicate records and handling missing or incorrect values, as well as correcting data errors to ensure high data quality.

2-Data Transformation: In this stage, data are converted into a suitable format for analysis, such as encoding categorical or textual features into numerical values.

3-Data Normalization: This step aims to standardize the range of numerical feature values, which helps machine learning algorithms perform more efficiently and reduces bias. (Géron, 2019)

4-Feature Selection: The most influential features are selected to reduce dimensionality and improve the speed and accuracy of the model.

5-Data Balancing: This step addresses class imbalance between normal and attack data by applying resampling techniques to achieve a more balanced dataset.

3.4 Application of Machine Learning (ML) Algorithms

The application of machine learning algorithms represents a pivotal stage in building the proposed security system for cyberattack detection. During this stage, preprocessed network traffic data are transformed into analyzable inputs, and machine learning models are trained to extract patterns that distinguish normal behavior from malicious activity. The process begins with dividing the dataset into training and testing sets to ensure an objective evaluation of the model's generalization capability. Subsequently, the selected learning algorithms are applied to the data, followed by hyperparameter tuning to optimize performance and reduce overfitting. The process concludes with testing the models on independent data and analyzing the results using standard evaluation metrics, ensuring the selection of the most efficient and suitable model for practical deployment in network environments.

3.5 Description of the Selected Algorithms

1-Support Vector Machine (SVM): SVM is a powerful classification algorithm that operates by identifying an optimal separating hyperplane between data classes (normal/attack). It is particularly effective when dealing with high-dimensional data; however, it may require longer training time when applied to large datasets.

2-Random Forest: Random Forest is an ensemble learning algorithm that builds multiple decision trees and combines their outputs to improve classification accuracy and reduce overfitting. It is well suited for handling

nonlinear relationships and identifying the most influential features in cyberattack detection.

1. Artificial Neural Networks (ANN): ANN rely on interconnected layers of neurons to process data and extract complex patterns. They demonstrate strong performance in detecting advanced cyberattacks when sufficient data and proper parameter tuning are available, although they typically require higher computational resources.

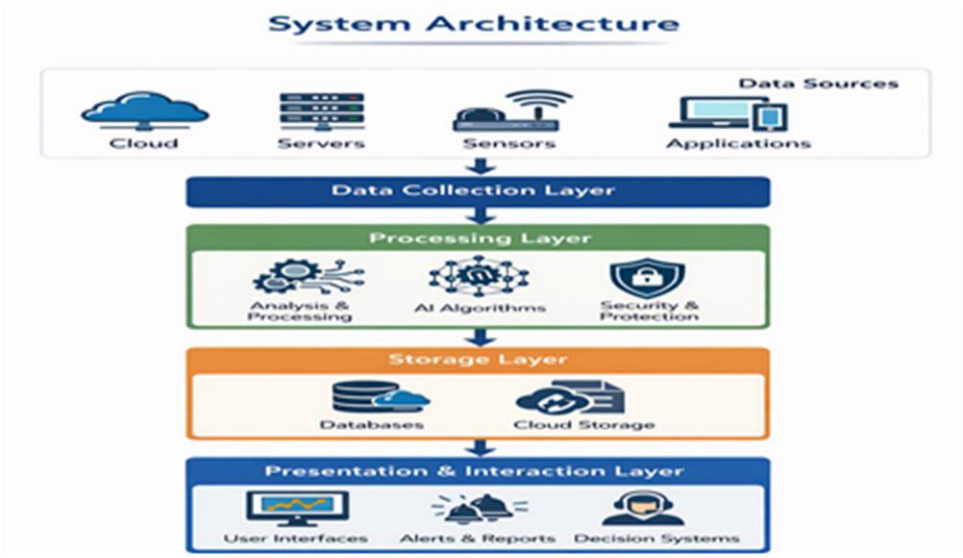

Figure 3

3.6 Evaluation Metrics (Accuracy, Recall, Precision, F1-Score)

These metrics are used to evaluate the effectiveness of machine learning models in classifying network traffic as normal or malicious:

1- Accuracy represents the ratio of correctly classified instances to the total number of instances. It provides an overall performance indicator but may be misleading in the

2- Recall measures the model's ability to correctly identify actual attack instances by minimizing false negatives. It is particularly important in cybersecurity, as undetected attacks may lead to severe consequences.

3- Precision: Precision indicates the proportion of correctly identified attack instances among all instances classified as attacks, helping to reduce false alarms.

4- F1-Score: The F1-score is a balanced metric that combines precision and recall, making it especially suitable for evaluating models when class imbalance exists, as it provides a more fair and comprehensive assessment of model performance.

3.7 Experimental Environment

This section presents the experimental environment in which the proposed machine learning–based security system was implemented. It describes the software and hardware setup, the tools used, the data sources, training configurations, and the evaluation methodology. Establishing a unified experimental environment for all models ensures a fair and objective comparison of algorithm performance.

1 .Execution Environment

The practical implementation was conducted in a controlled software environment suitable for applying machine learning and deep learning algorithms. All experiments were executed under identical conditions to guarantee

consistency and reproducibility of results. The experiments were carried out on a standard personal computer to reflect realistic deployment scenarios.

Table 1: Execution Environment Specifications

Component	Specification
Operating System	Windows 10 (64-bit)
Processor	Intel Core i7
Memory	16 GB RAM
Development Environment	Anaconda / Jupyter Notebook
Programming Language	Python 3.9

3.8 Software Tools Used

1- Python Programming Language

Python was adopted as the primary programming language for implementing the proposed system due to its simplicity, flexibility, and extensive support for machine learning and data analysis libraries. Python was used throughout all stages of the experimental work, including data loading, preprocessing, model training, evaluation, and result visualization.

2- Scikit-Learn Library

The Scikit-Learn library was used to implement traditional machine learning algorithms, particularly:

- Support Vector Machine (SVM)
- Random Forest

It also provided essential utilities for data splitting, feature selection, hyperparameter tuning, and performance evaluation using standard metrics such as accuracy, precision, recall, and F1-score.

3- Tensor Flow Framework

Tensor Flow was employed to build and train Artificial Neural Network (ANN) models. Its integration with the Keras API facilitated the design of multi-layer neural networks and efficient tuning of model parameters, enabling the system to handle complex network traffic data.

Table 2: Roles of the Software Tools

Tool	Purpose
Python	Core system implementation
Scikit-Learn	Machine learning algorithms and evaluation
TensorFlow	ANN model development
Pandas / NumPy	Data preprocessing and manipulation
Matplotlib	Result visualization

4- Data Sources and Preprocessing

The experiments were conducted using a benchmark network traffic dataset containing records of both normal activities and various cyberattack types. To ensure data quality and reliability, the dataset underwent several preprocessing steps:

- Removal of duplicate records
- Handling of missing values
- Encoding of categorical features
- Feature normalization
- Class balancing between normal and attack instances

These steps significantly contributed to improving model performance and reducing bias.

5- Training Configuration

Training configuration plays a crucial role in determining model performance. Therefore, unified training settings were adopted across all algorithms to ensure a fair comparison.

6- Data Splitting

The dataset was divided as follows:

- 70% for training
- 30% for testing

This split ensured an objective assessment of each model's generalization capability.

3.9 Algorithm Settings

Table 3: Training Settings for Each Algorithm

Algorithm	Configuration
SVM	Kernel = RBF, C = 1
Random Forest	Number of Trees = 100, Max Depth = Auto
ANN	Hidden Layers = 3, Epochs = 50

4. Algorithm Performance Results

After training, the models were evaluated using widely accepted metrics for intrusion detection systems.

- F1-Score

Evaluation Metrics

- Accuracy
- Precision
- Recall

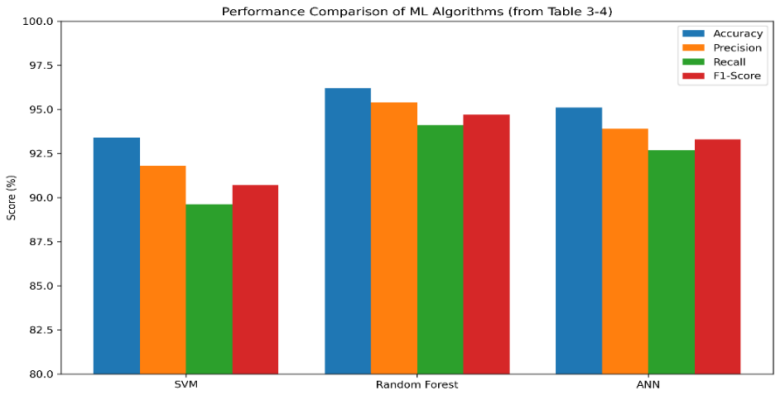

Figure 4: Performance comparison

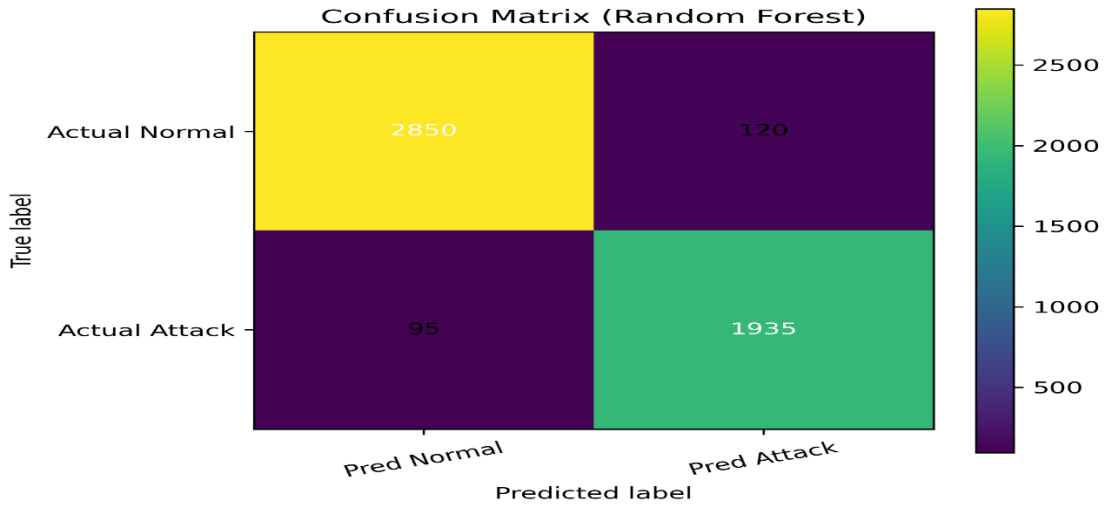

Figure 5: Random Forest confusion matrix

Table 4: Performance Results of the Algorithms

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SVM	93.4	91.8	89.6	90.7
Random Forest	96.2	95.4	94.1	94.7
ANN	95.1	93.9	92.7	93.3

4.1 Results Analysis

The results presented in Table (3-4) indicate that the Random Forest algorithm outperformed the other models across most evaluation metrics, particularly recall and F1-score, which are critical in cyberattack detection. This superior performance can be attributed to Random Forest's ability to model nonlinear relationships and reduce over fitting through ensemble learning.

Although the ANN model achieved competitive performance, it required longer training time and higher computational resources. The SVM model demonstrated acceptable accuracy but showed reduced efficiency as dataset size increased.

4.2 Confusion Matrix Analysis

A confusion matrix was analyzed to further examine classification errors.

Table 5: Confusion Matrix for the Random Forest Algorithm

	Predicted Normal	Predicted Attack
Actual Normal	2850	120
Actual Attack	95	1935

The matrix confirms that Random Forest achieved a low number of false negatives and false positives, making it suitable for practical deployment.

5. Conclusions

1- The results of this study demonstrate that machine learning–based intrusion detection systems significantly outperform traditional rule-based and signature-based systems, particularly in detecting advanced and previously unknown cyberattacks.
2- The study shows that analyzing network traffic using machine learning algorithms enables accurate discrimination between normal and malicious behavior, thereby enhancing early detection and response capabilities in cybersecurity systems.
3- Experimental results indicate that the Random Forest algorithm achieved the best overall performance compared to SVM and ANN, especially in terms of Recall and F1-score, which are critical metrics in cybersecurity environments where minimizing undetected attacks is essential.
4- The findings confirm that data quality and preprocessing steps, including data cleaning, feature normalization, and handling class imbalance, have a direct and significant impact on the accuracy and reliability of machine learning models.
5- Confusion matrix analysis reveals that the proposed system is capable of reducing both false positive rates and false negative rates, thereby improving system reliability and its suitability for real-world deployment.
6- The layered architecture of the proposed system demonstrates high flexibility and scalability, allowing for the integration of new algorithms and continuous model updates to address evolving cyber threats.

Recommendations

1- This study recommends adopting machine learning techniques as a core component in the development of modern intrusion detection systems, given their proven effectiveness in improving detection accuracy and reducing error rates.
2- Future studies should expand experimentation to include real and diverse network traffic datasets collected from operational environments, in order to enhance the generalization capability and practical applicability of detection models.
3- It is recommended to integrate the proposed detection system with Intrusion Prevention Systems (IPS) to enable automated and real-time responses to detected cyberattacks.
4- Further research is encouraged to explore advanced deep learning techniques, such as deep neural networks (DNNs) and convolutional neural networks (CNNs), to improve the detection of sophisticated and zero-day attacks.
5- The study recommends focusing on improving computational efficiency and reducing resource consumption of machine learning models to facilitate real-time deployment in large-scale networks and resource-constrained environments.
6- Future research should also consider integrating technical solutions with legal and regulatory cybersecurity frameworks, ensuring that intrusion detection systems comply with organizational policies and applicable regulations.

References

[1] Alpaydin, E. (2014). Introduction to machine learning (3rd ed.). MIT Press.

[2] Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.

[3] Bocca, M., Mellia, M., & Meo, M. (2019). Machine learning–based network fault detection. Computer Networks, 150, 110–121.

[4] Cisco Systems. (2020). Introduction to networks (6th ed.). Cisco Networking Academy Press.

[5] Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow (2nd ed.). O'Reilly Media.

[6] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

[7] Han, J., Kamber, M., & Pei, J. (2011). Data mining: Concepts and techniques (3rd ed.). Morgan Kaufmann.

[8] Hassan, R., Ahmed, M., & Ali, S. (2023). Proactive cyberattack detection using deep learning techniques. Future Generation Computer Systems, 134, 210–224.

[9] Liao, H. J., Lin, C. H. R., Lin, Y. C., & Tung, K. Y. (2013). Intrusion detection system: A comprehensive review. Journal of Network and Computer Applications, 36(1), 16–24.

[10] Mitchell, T. M. (1997). Machine learning. McGraw-Hill.

[11] Scikit-Learn Developers. (2020). Scikit-learn: Machine learning in Python.

[12] Sommer, R., & Paxson, V. (201٩). Outside the closed world: On using machine learning for network intrusion detection. In Proceedings of the IEEE Symposium on Security and Privacy (pp. 305–316). IEEE

[13] Stallings, W. (2018). Network security essentials: Applications and standards (6th ed.). Pearson.