International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 15 Issue 3, March-2026

AI-Based Threat Detection and Prevention System for Cyber Attacks

Mr. Kiran Devanand Ibitkar[1], Prof. Dr. Monika okade[2], Prof. Dr. Sunil Khatal[3]

[1]Student, Sharadchandra Pawar College of Engineering, Dumbarwadi, Otur, Pune. India
[2]Project Guide, Sharadchandra Pawar College of Engineering, Dumbarwadi, Otur, Pune. India
[3]HOD, Sharadchandra Pawar College of Engineering, Dumbarwadi, Otur, Pune. India

ABSTRACT

Cyber-attacks are becoming more common and more advanced. Because of this, older cyber security techniques, such as firewalls and antivirus software, are no longer enough to protect your network. To fix these issues, networks should have more advanced protection systems. One of these protective systems is Intrusion Detection Systems, or IDS. IDS can be either signature-based or anomaly-based. Signature-based systems are good at identifying previously known attacks, and systems built to identify new attacks will almost always fail. Anomaly-based systems can identify new types of attacks, but often will have to identify attacks that are not really happening, which is known as a false positive. To solve these problems, we created new machine learning and deep learning based network intrusion detection system for the cloud. This system will separate data into several parts and also will use techniques known as normalization and class balancing to improve system accuracy. Using the DeepLearning4J software, we tested several datasets and achieved 96.50% accuracy with the NSL-KDD datasets. We found that a system built with this type of machine learning is the better method for solving problems with network security.

Keywords— Machine Learning, Intrusion Detection System (IDS), Deep Learning, Cyber security Network Security.

INTRODUCTION

The focus here is the absence of human supervisors to enable autonomous threat identification. If a threat is detected, engaged, and if, automation is appropriately used elimination of the threat with little or no human assistance. If there is little or no human engagement in the process, the system is autonomous. Some human intervention may be used if automations fails or is inefficient. However, every instance of the human intervention is considered a defect or a failure of the system. The remote engagement, if used, is not designed for routine use. Autonomous reporting is possible, however the system may function without it. If the report is not automated, the system may function without it.

Using AI, Deep Learning, and Machine Learning, threat detection systems are able to analyze large-scale network and behavioral data to automatically recognize and respond to cyber-attacks on behavioral data to recognize and respond to cyber-attacks in real time [1]. Machine learning-based intrusion detection significantly improves accuracy and reduces false alarms, especially when handling imbalanced or complex network datasets [2]. AI and ML techniques strengthen cyber defense mechanisms by providing real-time detection and adaptive responses to new and evolving threats [4]. Swarm-optimized machine learning approaches have demonstrated improved threat prediction and classification performance in cybersecurity applications. ML-based anomaly detection methods are increasingly used in industrial control systems to identify intrusions and operational threats. Modern ICS environments rely on ML-driven defense strategies to secure system components and detect sophisticated cyberattacks [7] [9].

AI-enabled intelligent engineering systems enhance cyber threat detection by automating the recognition of abnormal activity patterns. Hybrid ML–DL models provide improved intrusion detection and prevention in Industry 4.0 wireless sensor networks. ML-integrated frameworks designed for smart grid systems offer advanced intrusion detection by combining traffic analysis and node categorization techniques [10]..

LITERATURE REVIEW

Kumar, Busireddy Hemanth, et al. (2025) [1] the shows how the fusion of artificial intelligence combined with machine learning and big data analytics enhances the identification of cyber threats. The authors highlight the unprecedented development of data in the field of cybersecurity and the challenges that its volume, velocity, and variety generate. They offer processing architectures that are scalable and that can facilitate the real-time detection of sophisticated attacks. Their results show that AI-based models are more effective than previous models based on rules in anomaly detection. Overall, the paper stresses the vital role of data-intensive intelligence in shaping future cybersecurity solutions.

Ahmed, Hafiza Anisa, et al. (2022) [2] An example of this is the researchers' design of an intrusion detection framework using machine learning with an emphasis on addressing dataset imbalance using oversampling techniques. They evaluate a handful of algorithms, such as Support Vector Machines and Random Forest, on established intrusion datasets. The balancing approach significantly boosts the classification of underrepresented attack categories. Experimental outcomes demonstrate better detection accuracy and fewer false negatives. The authors ultimately show that addressing data imbalance is crucial for dependable intrusion detection.

Deshmukh, et.al (2024) [3] an IDS built on convolutional neural networks, optimized for use in IoT environments where resources are limited. The CNN model analyzes compact network traffic features to identify malicious behaviors with strong accuracy. The authors illustrate that CNNs provide more effective feature extraction compared to traditional ML models. Tests conducted on IoT-specific datasets exhibit high performance. The study concludes that deep learning presents an efficient pathway for boosting IoT security.

Hussain, Hasnain, et al. (2025) [4] the work investigates how artificial intelligence and machine learning enhance modern cyber defense mechanisms. By contrasting classic signature-based detection with intelligent models capable of spotting unfamiliar threats, the authors display the limitations of traditional approaches. Their experiments highlight the importance of adaptive security systems in increasingly complex networks. Machine learning techniques demonstrate faster and more accurate threat identification. The research confirms AI’s pivotal role in addressing emerging cybersecurity challenges.

Qiqieh, Issa, et al. (2025) [5] the introduce a cyber threat detection framework that leverages swarm optimization to refine machine learning model parameters. Using metaheuristic methods such as Particle Swarm Optimization, they fine-tune classifiers to achieve greater accuracy and resilience when handling diverse cyberattacks. Large-scale datasets validate the improved performance of the optimized system. The outcomes show that swarm-based optimization delivers superior detection capability compared with traditional ML approaches.

Benka, Denis, et al. (2025) [6] centers on applying ML to detect anomalies and intrusions within Industrial Control Systems. The authors construct a specialized detection architecture tailored to the distinctive traffic behaviors and operational constraints of ICS environments. They test multiple ML models to assess real-time anomaly identification. Results indicate significant gains in detecting malicious actions and operational faults. The study contributes to enhancing the security posture of critical industrial infrastructures.

Nankya, Mary, et al. (2023) [7] the authors examine the fundamental elements of ICS ecosystems and outline prevalent cyber threats that endanger industrial processes. They suggest using ML-based defense methods to overcome these issues. The research demonstrates ML's potential to enhance the response times and resilience of industrial networks. Real-world case studies demonstrate the strategy’s effectiveness against operational attacks. The paper offers a broad overview of how ML strengthens ICS cybersecurity.

Sivakumar, Janaki, et al. (2025) [8] an AI-powered threat detection approach aimed at reinforcing the security of advanced engineering systems. The proposed architecture integrates multiple forms of AI to identify anomalies within intricate network environments. Emphasis is placed on building a system that is scalable, adaptive, and capable of automated responses. Experimental findings reveal improvements in detection precision and reductions in false alarms. The authors advocate integrating intelligent systems into modern cybersecurity frameworks.

Fatima, et al. (2024) [9] a predictive intrusion detection and prevention mechanism for Industry 4.0 wireless sensor networks. Deep learning and machine learning models are used to achieve higher accuracy when classifying network behaviors and detecting anomalies. It incorporates situational awareness to improve decision-making under changing network conditions. Experiments show high detection rates across multiple attack variations. The work supports hybrid AI strategies as effective solutions for securing industrial WSN environments.

Zhukabayeva, Tamara, et al. (2024) [10] introduces a machine learning–assisted intrusion detection and prevention framework for smart grid wireless sensor networks. The method integrates traffic profiling and node classification to heighten anomaly detection accuracy. It dynamically adapts to differing node roles and behavioral patterns within the grid. Experimental evaluation confirms enhanced threat detection and lower misclassification rates. The study strengthens cybersecurity approaches in energy-sector WSN applications.

METHODOLOGY

This study's foundation includes the NSL-KDD, which has multiple features of data that describe both benign and malicious activities. There are four steps involved in the method of using machine learning and deep learning for prevention and detection of cyber threats. They include data preprocessing, feature extraction, model building, and finally, evaluation of model performances. Data preprocessing involves imputing missing values using mean or median strategies, eliminating duplicate entries, and applying Min-Max normalization to scale numerical features. Subsequently, feature-importance analysis is performed to identify attributes that significantly influence intrusion detection accuracy.

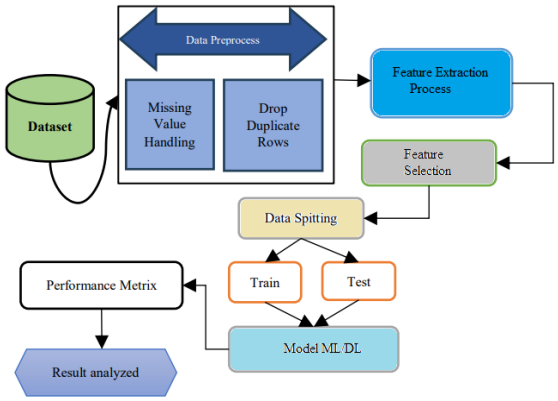

Fig.1 System Architecture

Dataset Description

The study utilized four widely used intrusion detection benchmarks: NSL-KDD, ISCX, BOTNET, and UNSW-NB15. NSL-KDD offers 41 structured attributes that represent basic connection details, content-based indicators, and various traffic statistics. ISCX provides realistic flow-oriented information, including metrics such as packet and byte counts, session duration, and packet inter-arrival intervals. The BOTNET dataset focuses on capturing malicious botnet activity through packet- and flow-level descriptors, including connection timing, packet size patterns, and command-and-control communication traces. UNSW-NB15 contributes 49 contemporary features produced through Argus and Bro/Zeek, spanning flow, basic, content, temporal, and additional traffic categories. Collectively, these datasets supply a broad and complementary range of features for assessing intrusion detection techniques across traditional, real-world, botnet-focused, and modern attack scenarios.

Data Preprocessing

Preprocessing the data is extremely important to ensure that the models outcome is as accurate as possible, while also making sure the dataset is reliable. In this stage, missing values are addressed and duplicate entries are removed to maintain data consistency. The steps involved are as follows:

- Handling Missing Values: Missing entries in any column are replaced using statistical measures such as the mean or median of that column, or substituted with a predefined constant value when appropriate.
- Removing Duplicate Rows: Duplicate records are identified and deleted from the dataset, ensuring the DataFrame contains only unique entries and eliminating unnecessary redundancy.

Feature Extraction

During the feature extraction phase, essential numerical and normal features are derived from the raw text data. Various techniques—such as correlation-based methods, co-occurrence analysis, relational feature extraction, and dependency-based features—are applied to obtain meaningful representations while eliminating irrelevant attributes from the dataset.

Feature Selection
After extracting the features, a refined subset is generated through feature selection, guided by specific quality thresholds. To optimize the feature set, a weighted term frequency approach is employed. The selected features are then passed to the training module to enhance model efficiency and performance.

Data Splitting
The data set was split into 80% training, 20% testing. This split lowers the chances of overfitting and gives a more reliable assessment of the model's performance on data it has not seen before. Keeping this ratio allows the model to be trained sufficiently and assessed sufficiently with the test set held out.

Classifier
At this stage, the dataset is divided into subclasses of attack or normal using some supervised learning methodologies. Also, the system can detect some novel (unknown) attack patterns embedded in normal-time data. In this research both machine learning and deep learning techniques are used in supervised classification. With the availability of some labeled data, the strategies of supervised machine learning algorithms namely Random Forest (RF) and the deep learning model Recurrent Neural Network (RNN) are implemented to detect identity deception in social networks. The RF classifier builds a number of decision trees based on randomly selected subsets of features, and the final class label is assigned based on the majority vote from all these trees. In addition, the RNN model improves the ability to make decisions based on sequences.

Performance Metrics
Performance metrics are employed to assess and compare algorithm performance using key indicators that reflect how effectively network intrusions are detected, with values generally ranging from 0 to 1. The confusion matrix's core elements—True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN)—are important for understanding the model's reliability and accuracy and for figuring out how well the system identifies normal and attack traffic.

Model performance is assessed using the confusion matrix, from which evaluation metrics such as F1-Score and PR-AUC are calculated to measure the effectiveness of the classification process

RESULTS AND DISCUSSION

To evaluate system performance, accuracy matrices were generated to measure the effectiveness of the implemented model. The system was developed in an open-source environment using a Java-based 3-tier architectural framework and executed on hardware consisting of an Intel 2.8 GHz i3 processor and 4 GB RAM. After implementation, a comparative assessment was conducted between the proposed system and several existing methods.

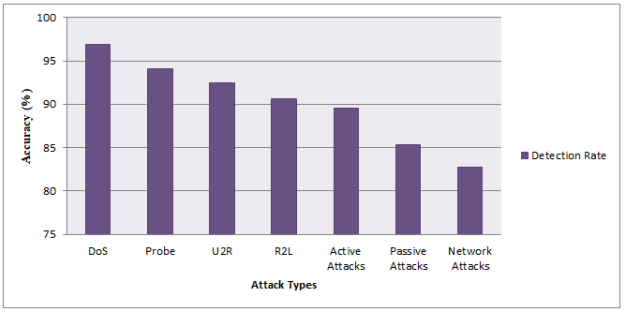

Fig. 2 Detection rate of proposed system for all attacks

Fig. 2 illustrates the average detection rate achieved for various attack categories. The system obtains a detection rate of 96.9% for DoS attacks, 94.10% for probing attacks, 92% for user-to-root attacks, and approximately 91% for remote-to-login attacks. In objective 1, the proposed algorithm achieves around 90% accuracy for active attacks, 85% accuracy for passive attacks, and an overall average accuracy of 83% for general network attacks.

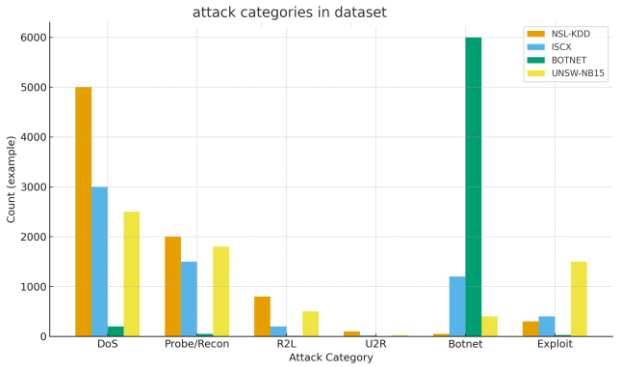

Fig. 3 Attack categories in dataset

The bar chart provides a comparative view of attack categories across the NSL-KDD, ISCX, BOTNET, and UNSW-NB15 datasets. It highlights how each dataset contains varying proportions of attacks such as DoS, Probe/Reconnaissance, R2L, U2R, Botnet, and Exploit. NSL-KDD features a high volume of DoS and Probe attacks, whereas ISCX presents a more evenly distributed set of attack types. The BOTNET dataset is predominantly composed of botnet-related traffic, evident from its significantly larger Botnet category. In contrast, UNSW-NB15 exhibits a diverse mix of attacks, with substantial representation of DoS, Probe, and Exploit events.

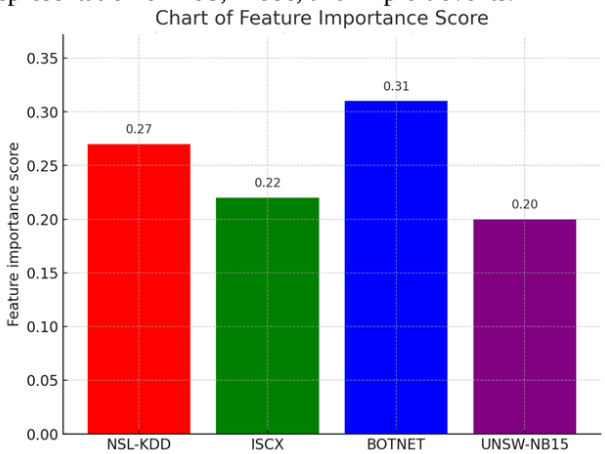

Fig. 4 Feature Importance Score

The bar chart presents a comparison of feature-importance scores across four widely used cybersecurity datasets: NSL-KDD, ISCX, BOTNET, and UNSW-NB15. Among these, the BOTNET dataset exhibits the highest feature importance, suggesting that its extracted features contribute most effectively to predictive accuracy. NSL-KDD also reflects strong feature relevance, while ISCX and UNSW-NB15 display moderate levels of influence.

CONCLUSION

This project created a new AI-based threat detection and prevention system that improves cyber security in different types of networks. By combining enhanced feature extraction methods with supervised machine learning and deep learning models,, the suggested system showed advanced capabilities in the detection of cyberattacks which are both eluded and disclosed. Detection compilation using the benchmarks of the multiple datasets of intrusion detection—NSL-KDD—demonstrated that the proposed system attained a level of effectiveness and a level of detection consistency within different classes of attacks. Using machine learning and deep learning to analyze patterns in network traffic, system logs, and user behavior, the system identifies anomalous and emerging threats. Its automated preventive capabilities enhance response times and decrease the impact of attacks. Additionally, the system improves from experience at dealing with new cyber threats. Because of its real-time capabilities, the system creates a secure environment that allows business owners to act quickly and with confidence. This AI-based system creates a reliable, scalable, and smart foundation that offers complete cybersecurity. It is a crucial resource for the protection of critical infrastructure and sensitive data in today's cyber environment.

REFERENCES

[1] Kumar, Busireddy Hemanth, et al. "Big Data in Cybersecurity: Enhancing Threat Detection with AI and ML." Metallurgical and Materials Engineering 31.3 (2025): 12-20.
[2] Ahmed, Hafiza Anisa, Anum Hameed, and Narmeen Zakaria Bawany. "Network intrusion detection using oversampling technique and machine learning algorithms." PeerJ Computer Science 8 (2022): e820.
[3] Deshmukh, Amogh, and Kiran Ravulakollu. "An efficient CNN-based intrusion detection system for IoT: Use case towards cybersecurity." Technologies 12.10 (2024): 203.
[4] Hussain, Hasnain, Maria Kainat, and Taib Ali. "Leveraging AI and machine learning to detect and prevent cyber security threats." Dialogue Social Science Review (DSSR) 3.1 (2025): 881-895.
[5] Qiqieh, Issa, et al. "An intelligent cyber threat detection: A swarm-optimized machine learning approach." Alexandria Engineering Journal 115 (2025): 553-563.
[6] Benka, Denis, et al. "Machine Learning-Based Detection of Anomalies, Intrusions and Threats in Industrial Control Systems." IEEE Access (2025).
[7] Nankya, Mary, Robin Chataut, and Robert Akl. "Securing industrial control systems: Components, cyber threats, and machine learning-driven defense strategies." Sensors 23.21 (2023): 8840.
[8] Sivakumar, Janaki, et al. "AI-driven cyber threat detection: enhancing security through intelligent engineering systems." Journal of Information Systems Engineering and Management 10.19 (2025): 790-798.
[9] Al-Quayed, Fatima, Zulfiqar Ahmad, and Mamoona Humayun. "A situation based predictive approach for cybersecurity intrusion detection and prevention using machine learning and deep learning algorithms in wireless sensor networks of industry 4.0." Ieee Access 12 (2024): 34800-34819.
[10] Zhukabayeva, Tamara, et al. "A traffic analysis and node categorization-aware machine learning-integrated framework for cybersecurity intrusion detection and prevention of WSNs in smart grids." IEEE Access 12 (2024): 91715-91733.