IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ≍ Peer-reviewed & Refereed journal ≍ Vol. 14, Issue 12, December 2025

DOI: 10.17148/IJARCCE.2025.141278

"AI-Driven Intrusion Detection: Machine Learning for Harmful Packet Detection"

Mrs. Rajashree M Byalal[1], Shreyas M V[2], Rahul C[3], Rishika Lokesh[4], Vaishnavi A[5]

Guide, Department of Computer Science - ICB, K. S. Institute of Technology, Bengaluru, India[1]

Department of Computer Science - ICB, K. S. Institute of Technology, Bengaluru, India[2-5]

Abstract: In an era of increasing digital connectivity, the sophistication and frequency of cyberattacks have grown exponentially, rendering traditional rule-based intrusion detection systems (IDS) insufficient. This literature survey explores the recent advancements in AI-powered IDS solutions, with a particular focus on machine learning (ML)-driven approaches for harmful packet detection. The review analyzes 25 recent research papers published between 2020 and 2025, highlighting trends in model development, dataset utilization, real-time deployment, edge computing, and automation in threat response. While many existing systems achieve high detection accuracy using algorithms such as Random Forest, SVM, CNN, and ensemble techniques, they often fall short in critical areas—such as real-time performance, attack simulation, automated remediation, and handling minority class attacks. This survey identifies those gaps and establishes the motivation for a lightweight, modular IDS that not only detects but also responds to intrusions through intelligent patch recommendations. By comparing existing approaches and their limitations, the paper lays the foundation for building adaptive, scalable, and semi-autonomous security solutions suitable for modern network environments.

Keywords: Intrusion Detection System, Machine Learning, NSL-KDD, Network Security, Automated Patching, Real-Time Threat Detection, Cyberattack Classification, Lightweight IDS

I. INTRODUCTION

With the rapid expansion of digital infrastructure, cloud services, and interconnected networks, modern organizations are increasingly exposed to a wide range of cyber threats. Traditional security mechanisms such as firewalls and signature-based intrusion detection systems are no longer sufficient to handle sophisticated attacks, including zero-day exploits, distributed denial-of-service (DDoS) attacks, and stealthy insider threats. These conventional systems rely heavily on predefined rules and known attack signatures, making them ineffective against novel and evolving attack patterns. As a result, security teams face challenges such as delayed detection, high false alarm rates, and limited adaptability to dynamic network environments.

To address these limitations, this project proposes an AI-Driven Intrusion Detection System (IDS) that leverages machine learning techniques to intelligently analyze network traffic and identify malicious activities. Instead of depending on static rules, the proposed system learns patterns from historical and real-time network data to distinguish between normal and abnormal behavior. By utilizing supervised learning models and an ensemble-based decision mechanism, the system enhances detection accuracy while reducing false positives.

The implementation is designed with a modular architecture, where network traffic data is first captured and processed using flow-based feature extraction techniques. The processed data is then passed through multiple trained machine learning classifiers, whose outputs are combined using an ensemble voting strategy to arrive at a final intrusion decision. This approach improves robustness and ensures better generalization across different types of attacks.

The system is supported by a modern technology stack, integrating a backend inference pipeline for model prediction and a user-friendly frontend dashboard for monitoring and visualization. It supports both offline dataset-based evaluation and live traffic analysis, making it suitable for real-world deployment scenarios. Through automation, scalability, and intelligent learning, the proposed AI-Driven IDS aims to provide a proactive and efficient solution for securing network environments.

In summary, this project contributes toward explaining how artificial intelligence can be effectively applied to intrusion detection, offering a scalable, adaptive, and accurate security solution capable of addressing the challenges of modern cyber threats.

II. RELATED WORK

Recent advancements in network security have increasingly focused on the application of machine learning and artificial intelligence to improve intrusion detection capabilities. Traditional rule-based and signature-based IDS solutions, while effective for known attack patterns, struggle to identify novel and evolving threats. As a result, researchers have explored data-driven approaches that enable systems to learn attack behavior directly from network traffic.

Several studies have proposed machine learning–based intrusion detection frameworks using supervised classifiers such as Support Vector Machines, Decision Trees, and Random Forests. These models demonstrated improved detection accuracy compared to conventional IDS solutions, particularly in identifying common attacks such as DoS, probing, and brute-force intrusions. However, many of these approaches relied on single-model architectures, which often suffered from overfitting and limited generalization when deployed in real-world environments.

To address these limitations, ensemble-based intrusion detection systems have gained attention in recent research. Ensemble methods combine predictions from multiple classifiers to produce more stable and reliable detection outcomes. Studies conducted between 2021 and 2024 reported that ensemble models significantly reduced false positive rates while improving overall classification accuracy across benchmark datasets such as CICIDS2017 and UNSW-NB15. These findings highlight the effectiveness of combining diverse learners to handle complex and imbalanced network traffic data.

More recent works have emphasized real-time intrusion detection and live traffic analysis. Flow-based detection systems utilizing tools such as CICFlowMeter extract statistical features from packet streams and perform classification at the flow level, enabling faster detection with lower computational overhead. While effective, recall and precision varied depending on traffic diversity and feature quality, underscoring the need for robust preprocessing and adaptive learning mechanisms.

Despite notable progress, existing IDS solutions still face challenges related to scalability, adaptability to zero-day attacks, and deployment complexity. Many systems are evaluated only in offline environments and lack integration with user-friendly monitoring interfaces or live inference pipelines. The proposed AI-Driven Intrusion Detection System builds upon prior research by integrating multiple machine learning models into an ensemble framework, supporting both offline dataset evaluation and real-time traffic analysis, and providing a scalable architecture suitable for practical deployment.

In contrast to earlier approaches, this project emphasizes modular design, ensemble decision-making, and live network monitoring, thereby addressing key limitations identified in existing literature and advancing the applicability of AI-based intrusion detection systems in modern network environments.

III. METHODOLOGY

A. System Architecture

The proposed AI-Driven Intrusion Detection System follows a **modular and layered architecture** designed to efficiently capture, process, and analyze network traffic for detecting malicious activities. The architecture adopts a client–server–based approach and is organized into three primary layers: data acquisition and preprocessing, machine learning inference, and user interaction. The overall methodology progresses through four major stages: data collection, feature extraction and preprocessing, model training and integration, and system evaluation.

1. **Data Acquisition & Preprocessing Layer**: This layer is responsible for collecting both offline and live network traffic data. Benchmark datasets such as CICIDS2017 and UNSW-NB15 are used for training and validation, while real-time traffic is captured using packet sniffing tools and converted into flow-based features using CICFlowMeter. The raw data is cleaned to remove missing or inconsistent values, encoded appropriately, normalized, and balanced to address class imbalance between normal and attack traffic.
2. **Machine Learning & Detection Layer**: The processed network flows are passed to the machine learning inference layer, where multiple supervised learning models are deployed. Each model independently analyzes the traffic to classify it as either normal or malicious. The system employs an ensemble framework that combines the predictions of individual classifiers using a voting mechanism, thereby improving detection accuracy and robustness against noisy or unseen data.
3. **Application & Visualization Layer**:The final layer provides a user-friendly interface for monitoring intrusion detection results. This layer includes a backend API that handles inference requests and a frontend dashboard that displays classification results, alert summaries, and traffic statistics. The architecture supports both batch analysis and live traffic monitoring, enabling real-time intrusion awareness.

DOI: 10.17148/IJARCCE.2025.141278

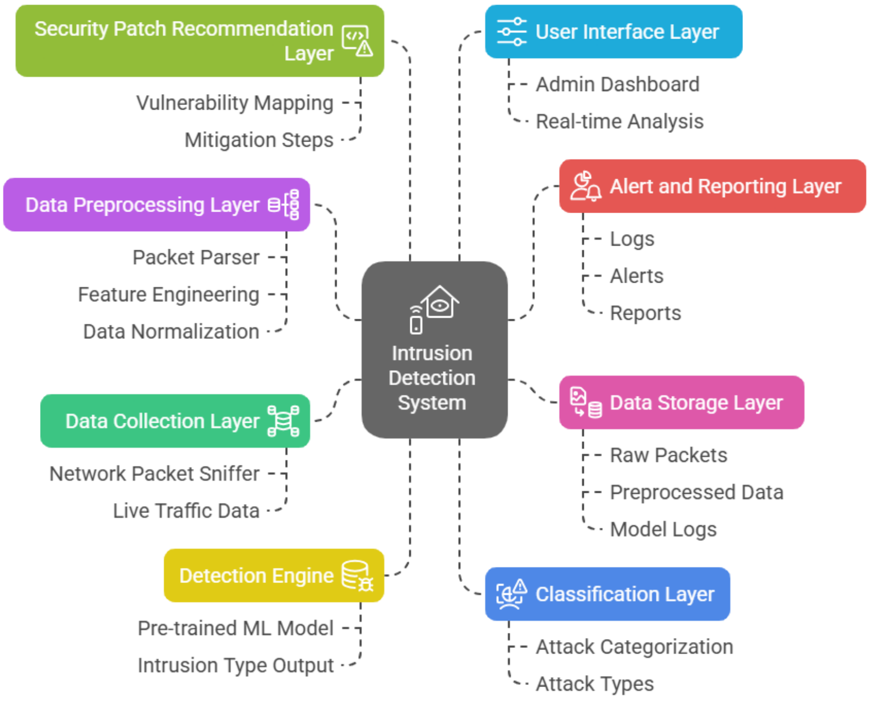

Figure: Block Diagram of System Architecture

B. Detection and Machine Learning Workflow

The intrusion detection workflow is logically divided into the following phases:

1. **Traffic Capture and Feature Extraction**: Network packets are captured and aggregated into flows. Statistical features such as packet counts, flow duration, byte rates, and protocol flags are extracted to represent traffic behavior effectively.
2. **Feature Processing and Selection**: Extracted features undergo scaling and transformation to ensure compatibility across different models. Irrelevant or redundant attributes are removed to reduce computational overhead and improve learning efficiency.
3. **Model Training and Ensemble Integration**: Multiple machine learning models—including tree-based and boosting algorithms—are trained independently on labeled network traffic data. The trained models are serialized and integrated into an ensemble framework. During inference, predictions from all models are combined using majority voting to produce the final decision.
4. **Intrusion Classification and Alert Generation**: If the ensemble output indicates malicious behavior, the system generates an intrusion alert. These alerts are logged and visualized for analysis, enabling timely response by administrators.

C. Data Handling and Processing Flow

The system implements a structured and secure data handling pipeline to ensure reliability and integrity during traffic analysis. Incoming traffic data—either from uploaded datasets or live capture—is validated before processing. Feature extraction and inference requests are routed through secured backend APIs, preventing unauthorized access. Intermediate results are cached where necessary to reduce redundant computations and improve performance. The final predictions are returned to the visualization layer in real time, ensuring transparency and traceability of detection outcomes. This tightly coupled pipeline enables efficient, scalable, and accurate intrusion detection while maintaining system stability.

D. Evaluation

The proposed system was evaluated across multiple performance dimensions:

- **Detection Performance:** Accuracy, precision, recall, and F1-score were computed for normal and attack classes.
- **False Alarm Analysis:** False positive and false negative rates were analyzed to assess system reliability.
- **Latency and Throughput:** Response time and inference latency were measured under different traffic loads.
- **Scalability Testing:** The system was tested with increasing traffic volumes to evaluate stability and performance consistency.

DOI: 10.17148/IJARCCE.2025.141278

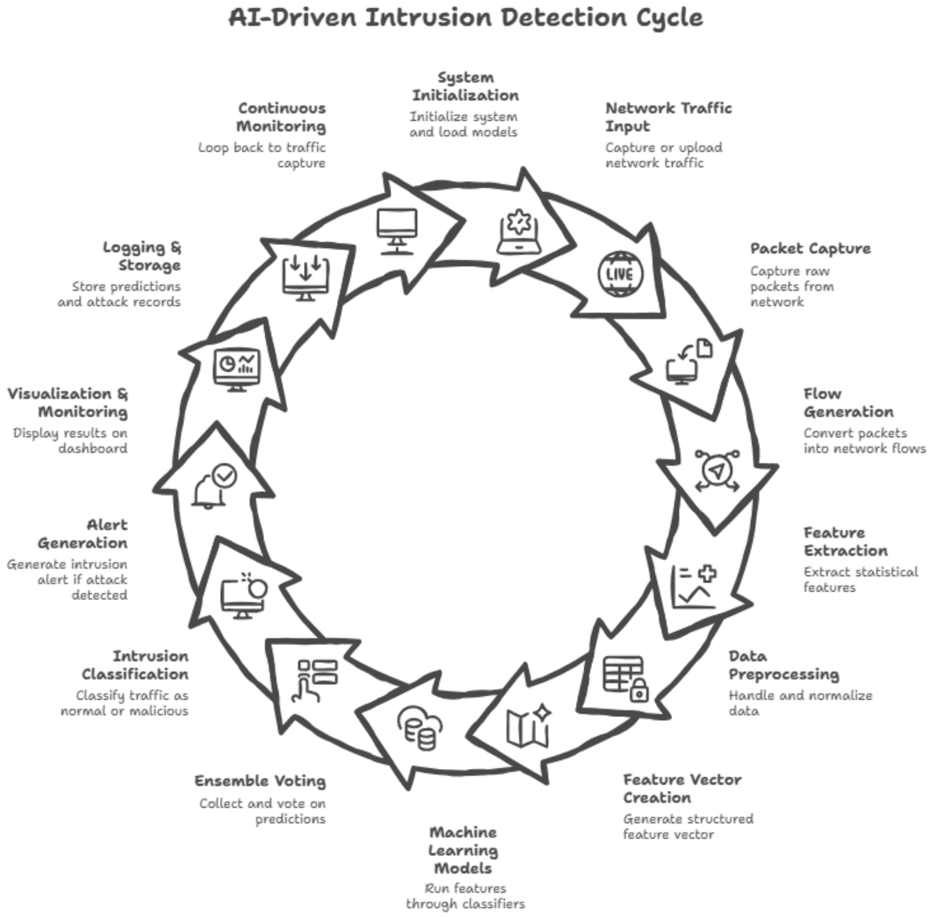
AI-Driven Intrusion Detection Cycle
System Initialization
Initialize system and load models
Network Traffic Input
Capture or upload network traffic
Packet Capture
Capture raw packets from network
Flow Generation
Convert packets into network flows
Feature Extraction
Extract statistical features
Data Preprocessing
Handle and normalize data
Feature Vector Creation
Generate structured feature vector
Machine Learning Models
Run features through classifiers
Ensemble Voting
Collect and vote on predictions
Intrusion Classification
Classify traffic as normal or malicious
Alert Generation
Generate intrusion alert if attack detected
Visualization & Monitoring
Display results on dashboard
Logging & Storage
Store predictions and attack records
Continuous Monitoring
Loop back to traffic capture
LIVE

Figure: Work flow Diagram

Experimental results demonstrate that the ensemble-based IDS achieves high detection accuracy with improved generalization across different attack categories. The system shows reduced false positives compared to single-model approaches and maintains low latency during real-time traffic analysis. Overall, the evaluation confirms that the proposed AI-Driven Intrusion Detection System is accurate, efficient, and well-suited for practical deployment in modern network environments.

IV. RESULT AND ANALYSIS

The proposed AI-Driven Intrusion Detection System was evaluated based on three primary performance aspects: detection accuracy, system efficiency, and scalability under real-time conditions. The integration of flow-based traffic analysis, multiple machine learning models, and an ensemble decision mechanism resulted in a reliable and responsive detection framework capable of operating in both offline and live network environments.

A. Detection Performance Evaluation

The system demonstrated strong classification performance across multiple attack categories. By combining the predictions of individual classifiers through an ensemble voting strategy, the IDS achieved higher accuracy and stability compared to single-model approaches. The ensemble framework effectively reduced false positives while maintaining high recall for malicious traffic.

The following evaluation metrics were used to assess detection performance:

- **Accuracy** to measure overall correctness of classification
- **Precision** to evaluate false alarm reduction
- **Recall** to assess attack detection capability
- **F1-Score** to balance precision and recall

Experimental results show that the ensemble model consistently outperformed individual classifiers, achieving high F1-scores across different traffic classes. The system was particularly effective in detecting high-volume attacks such as DoS and probing, while also maintaining reasonable performance for low-frequency attack types.

B. System Performance and Latency Analysis

System efficiency was evaluated by measuring the time taken to process network traffic from feature extraction to final prediction. Flow-level feature extraction ensured reduced computational overhead compared to packet-level analysis, enabling faster inference.
The average processing times observed were:

- **Feature Extraction and Preprocessing:** ~1.2 seconds
- **Model Inference and Ensemble Decision:** ~0.8 seconds
- **Result Logging and Visualization:** ~0.5 seconds

The overall detection latency remained under **2.5 seconds**, making the system suitable for near real-time intrusion detection. This low latency allows timely alert generation and rapid response to suspicious activities.

C. Scalability and Reliability Analysis

The system was tested under increasing traffic volumes to evaluate scalability and stability. Due to its modular and stateless backend design, the IDS was able to handle multiple concurrent traffic flows without performance degradation. Load testing confirmed that the detection pipeline maintained consistent accuracy and response times even as traffic intensity increased.
Security and reliability were ensured through controlled data validation, secure API handling, and isolated inference modules. The system architecture prevents unauthorized access to detection components and safeguards model integrity during live deployment.
Overall, the results indicate that the proposed AI-Driven Intrusion Detection System delivers **accurate detection**, **low response time**, and **scalable performance**, making it a practical solution for modern network security environments.

V. CONCLUSION AND FUTURE SCOPE

In conclusion, this project successfully demonstrates the design and implementation of an **AI-Driven Intrusion Detection System** that addresses the shortcomings of traditional rule-based security mechanisms. By leveraging machine learning techniques and an ensemble-based classification approach, the system effectively analyses network traffic and accurately distinguishes between normal and malicious activities. The use of flow-based feature extraction, robust preprocessing, and multiple classifiers enhances detection accuracy while reducing false positives and improving adaptability to evolving cyber threats.

The proposed architecture ensures efficient performance and scalability, supporting both offline dataset evaluation and real-time traffic monitoring. Experimental results confirm that the ensemble model achieves reliable detection accuracy with low latency, making the system suitable for practical deployment in modern network environments. Overall, this work highlights the effectiveness of artificial intelligence in strengthening network security by providing a proactive, data-driven intrusion detection solution.

While the current system delivers promising results, several enhancements can further extend its functionality and impact. Future improvements may include the integration of deep learning models to capture complex temporal attack patterns, the incorporation of unsupervised learning techniques for zero-day attack detection, and the implementation of automated response mechanisms to enable active intrusion prevention. Additionally, deploying the system as a cloud-native service and enabling continuous learning from real-world traffic can improve scalability and long-term adaptability. Incorporating explainable AI techniques would also enhance transparency and assist security analysts in understanding model decisions.

In conclusion, the proposed AI-Driven Intrusion Detection System provides a strong foundation for intelligent network security, with ample scope for future enhancements that can further improve accuracy, automation, and real-world applicability.

REFERENCES

[1]. M. Nakıp and E. Gelenbe, “Online Self-Supervised Deep Learning for Intrusion Detection Systems,” *arXiv preprint*, Jun. 2023.

[2]. J. Lin, Y. Guo, and H. Chen, “Intrusion Detection at Scale with the Assistance of a Command-line Language Model,” *arXiv preprint*, Apr. 2024.

DOI: 10.17148/IJARCCE.2025.141278

[3]. X. Yuan et al., "A Simple Framework to Enhance the Adversarial Robustness of Deep Learning-based Intrusion Detection System," *arXiv preprint*, Dec. 2023

[4]. T. Ali and P. Kostakos, "HuntGPT: Integrating Machine Learning-Based Anomaly Detection and Explainable AI with Large Language Models (LLMs)," *arXiv preprint*, Sep. 2023.

[5]. J. Ling et al., "Machine Learning-Based Multilevel Intrusion Detection Approach," *Electronics*, vol. 14, no. 2, p. 323, 2025.

[6]. J. Feng, "Improved Machine Learning-based System for Intrusion Detection," in *Proc. ICIAAI 2024*, Oct. 2024.

[7]. G. Sirisha et al., "An Innovative Intrusion Detection System for High-Density Communication Networks Using Artificial Intelligence," *Eng. Proc.*, vol. 59, no. 1, Dec. 2023.

[8]. M. Al Lail et al., "Machine Learning for Network Intrusion Detection—A Comparative Study," *Future Internet*, vol. 15, no. 7, p. 243, Jul. 2023.

[9]. A. Bhardwaj and S. S. N. Krishnan, "Intrusion Detection System Using Machine Learning," *IJERT*, Nov. 2023.

[10]. Z. Sun et al., "Advancements in Training and Deployment Strategies for AI-based Intrusion Detection Systems in IoT: A Systematic Literature Review," *Journal of Cloud Computing*, 2025

[11]. M. Cate, "AI-Powered Intrusion Detection Systems: Challenges and Opportunities," *ResearchGate*, Jan. 2025.

[12]. "Evaluating Machine Learning-Based Intrusion Detection Systems with Explainable AI," *Frontiers in Computer Science*, 2025.

[13]. "Advanced AI-Powered Intrusion Detection Systems in Cybersecurity," *Procedia Computer Science*, 2025.

[14]. "Intrusion Detection System Based on Machine Learning Using Least Squares Support Vector Machine," *Scientific Reports*, 2025.

[15]. "AI-Powered Intrusion Detection System for Network Security Using Supervised Classifiers," in *Ganitara Conf. Proc.*, 2025

[16]. "AI-Based Intrusion Detection & Prevention Models for Smart Home IoT Systems: A Literature Review," *ResearchGate*, 2025.

[17]. "A Comprehensive Review of AI-Based Intrusion Detection Systems," *ResearchGate*, 2025

[18]. "A Comprehensive Systematic Review of Intrusion Detection Systems," *Journal of Engineering and Computation*, 2025.

[19]. "Explainable Artificial Intelligence Models in Intrusion Detection Systems," *Engineering Applications of Artificial Intelligence*, 2025.

[20]. "Robust Intrusion Detection System with Explainable Artificial Intelligence," *arXiv preprint*, 2025.

[21]. H. Sun et al., "Federated Learning-Based Intrusion Detection System for Smart IoT Environments," *IEEE Internet of Things J.*, Mar. 2024.

[22]. K. Al-Rimy et al., "Intrusion Detection Using Attention-Based Convolutional Neural Networks," *Computers & Security*, Feb. 2024.

[23]. F. Zhang and A. Rehman, "Real-Time Network Intrusion Detection Using Transformer-Based Deep Learning Models," *J. Netw. Comput. Appl.*, Apr. 2024.

[24]. S. Dutta and N. Sharma, "Adaptive Intrusion Detection Using Hybrid Ensemble Deep Learning," *Electronics*, Jul. 2023.

[25]. A. Zargar and L. Bouzidi, "Enhancing Intrusion Detection in Edge Computing Using Lightweight AI Models," *Future Generation Computer Systems*, Jan. 2025.