



PDF Download
3787224.pdf
26 March 2026
Total Citations: 0
Total Downloads: 164

Latest updates: <https://dl.acm.org/doi/10.1145/3787224>

RESEARCH-ARTICLE

Aegis-5: A Hybrid Ensemble Framework for Intrusion Detection in Industry 5.0 Driven Smart Manufacturing Environment

VIJAY GOVINDARAJAN, Colorado State University, Fort Collins, CO, United States

FARAZ AHMED

ZAID BIN FAHEEM, Wuhan University, Wuhan, Hubei, China

MUHAMMAD BILAL, Rawalpindi Women University, Rawalpindi, Punjab, Pakistan

MANEL AYADI, Princess Nourah Bint Abdulrahman University, Riyadh, Ar Riyad, Saudi Arabia

JEHAD ALI, Ajou University, Suwon, Gyeonggi-do, South Korea

Open Access Support provided by:

Wuhan University

Ajou University

Rawalpindi Women University

Princess Nourah Bint Abdulrahman University

Colorado State University

Published: 21 January 2026
Accepted: 20 December 2025
Revised: 15 September 2025
Received: 12 May 2025

[Citation in BibTeX format](#)

Aegis-5: A Hybrid Ensemble Framework for Intrusion Detection in Industry 5.0 Driven Smart Manufacturing Environment

VIJAY GOVINDARAJAN, Colorado State University Department of Computer Information Systems 501 W Laurel St, Fort Collins, CO 80523, United State.

FARAZ AHMED, Crisp Technologies LLC Cybersecurity researcher, United State.

ZAID BIN FAHEEM, Department of Computer Science, Wuhan University, Wuhan, 430000, China.

MUHAMMAD BILAL, Department of Information Technology, Rawalpindi Women University, Pakistan.

MANEL AYADI, Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

JEHAD ALI, Department of AI Convergence Network, Ajou University, Suwon, South Korea.

Industry 5.0 represents a transformative paradigm that emphasizes synergy between human expertise, intelligent systems, and hyper connected cyber-physical environments. While this evolution fosters personalized automation and resilient production, it also amplifies the cybersecurity risks inherent in Industrial Internet of Things (IIoT) infrastructures. In this research, we present Aegis-5 a novel adaptive hybrid ensemble framework explicitly designed for intrusion detection in Industry 5.0-enabled smart manufacturing ecosystems. The proposed model integrates five diverse classifiers Random Forest, Gradient Boosting, XGBoost, SVM, and K-Nearest Neighbors using a dynamic weighting strategy guided by per-class precision, recall, and F1-score performance in real time. A meta-learner further synthesizes these predictions to enhance robustness against sophisticated and zero-day attacks. To ensure relevance and reliability, we evaluate the model using two benchmark IIoT datasets: IoT-23 and CIC-IoT 2023, both of which capture a broad spectrum of real-world industrial threats. Experimental results demonstrate that our framework achieves superior performance, with accuracy rates of 99.98% on IoT-23 and 99.95% on CIC-IoT 2023, coupled with precision (99.97%, 99.93%), recall (99.96%, 99.92%), and F1-score (99.96%, 99.93%) respectively., significantly reduces false positives, and adapts effectively to evolving attack behaviors. By aligning intelligent anomaly detection with the responsiveness and adaptability required by Industry 5.0, Aegis-5 offers a scalable, real-time, and practical cybersecurity solution for next-generation industrial systems.

CCS Concepts: • **Security and privacy** → **Artificial immune systems**.

Additional Key Words and Phrases: Industry 5.0, Smart Manufacturing, Hybrid Ensemble Framework, Real-Time Threat Detection, IIoT Security, Cyber-Physical Systems, Adversarial Training, Meta-Learning

*Corresponding author: jehadali@ajou.ac.kr

Authors' Contact Information: Vijay Govindarajan, Vijay.Govindarajan15@alumni.colostate.edu, Colorado State University Department of Computer Information Systems 501 W Laurel St, Fort Collins, CO 80523, United State.; Faraz Ahmed, Crisp Technologies LLC Cybersecurity researcher, United State.; Zaid Bin Faheem, Department of Computer Science, Wuhan University, Wuhan, 430000, Wuhan, 430000, China.; Muhammad bilal, Department of Information Technology, Rawalpindi Women University, Rawalpindi, Punjab, Pakistan.; Manel Ayadi, Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia; Jehad Ali, Department of AI Convergence Network, Ajou University, Suwon , Suwon, South Korea.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, or post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2026 Copyright held by the owner/author(s).

ACM 1556-4703/2026/1-ART

<https://doi.org/10.1145/3787224>

1 Introduction

The development of Industry 5.0 is a deeply transformative shift in manufacturing and industrial automation, marked by the transparent integration of human intelligence with sophisticated artificial intelligence (AI) and machine learning (ML) technology. Personalized production, human-centered collaboration, and smart decision-making within highly networked cyber-physical spaces are the hallmarks of this new era [28]. With industry advancement towards smart manufacturing, the role of strong cybersecurity structures is more important than ever. The convergence of Industrial Internet of Things (IIoT) devices, the backbone of Industry 5.0, results in dynamic and intricate systems that are susceptible to a variety of advanced cyber threats.

Cyber-attacks on these IIoT systems are highly risky [11], from data breaches to extreme industrial process disruptions. Legacy IDS and security products, built for static, traditional IT environments, are not effective to counter the special challenges presented by Industry 5.0. These challenges are the ultra-dynamic nature of IIoT traffic, device and communication protocol heterogeneity, and growing complexity in attack methods. Signature-based detection techniques and isolated machine learning models tend to be poorly suited to identify zero-day attacks or keep up with the changing threat environment, resulting in high false positive rates and low detection accuracy.

One of the solutions to these issues is the application of hybrid ensemble learning techniques, which leverage the strengths of multiple classifiers to enhance detection performance and responsiveness [26]. In contrast to conventional models, ensemble approaches take advantage of the diversity of different machine learning algorithms, e.g., Random Forest, Gradient Boosting, and XGBoost, to identify anomalies better. Nevertheless, even ensemble approaches may fail in the Industry 5.0 context, where the type of attacks can evolve quickly and in an unpredictable manner, requiring a more adaptive strategy. In this research, we present an Aegis-5 (Adaptive Ensemble for Guarding Industrial Systems in Industry 5.0) new ensemble-based system specifically tailored to meet the cybersecurity requirements of Industry 5.0 settings. The system employs a dynamic weighting scheme that adjusts the contribution of each classifier depending on real-time performance metrics such as precision, recall, and F1-score. This enables the system to focus on the most efficient models at various stages of attack detection. By combining classifiers like Random Forest, Gradient Boosting, XGBoost, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN), we develop a system that responds to known as well as unknown attack patterns in real-time. To ensure the efficacy of this method, we test the model on two leading IIoT datasets: IoT-23 and CIC-IoT 2023. These datasets represent a rich view of actual industrial network traffic and encompass a broad variety of attack scenarios, ranging from botnets and DDoS attacks to more advanced, targeted attacks. The outcomes show that our ensemble model delivers excellent performance, with accuracy rates of 99.98% on the IoT-23 dataset and 99.94% on the CIC-IoT 2023 dataset, lowering false positive rates significantly and enhancing overall detection ability. The primary contributions of this research are as follows:

- Proposing an Aegis-5 (Adaptive Ensemble for Guarding Industrial Systems in Industry 5.0) Dynamic Hybrid Ensemble Framework explicitly designed for intrusion detection in Industry 5.0-driven smart manufacturing ecosystems. The framework integrates five diverse classifiers—Random Forest, Gradient Boosting, XGBoost, SVM, and K-Nearest Neighbors—with a dynamic weighting strategy that adjusts model contributions in real time based on per-class precision, recall, and F1-score performance. This ensures adaptability to evolving and heterogeneous cyber threats.
- Enhancing Robustness via Meta-Learning and Hybrid Voting by employing a Logistic Regression meta-learner to synthesize base classifier predictions and combining soft and hard voting mechanisms. This hybrid approach mitigates bias, improves consensus, and strengthens detection accuracy against sophisticated and zero-day attacks.
- Leveraging Real-World IIoT Datasets (IoT-23 and CIC-IoT 2023) to validate the framework. These datasets capture diverse attack scenarios, including botnets, DDoS, and stealthy intrusions, ensuring relevance to

Industry 5.0 environments. Advanced preprocessing techniques, such as feature selection, subsampling, and protocol-specific engineering, optimize data quality and model generalizability.

- Achieving State-of-the-Art Performance with detection accuracy rates of 99.98% (IoT-23) and 99.94% (CIC-IoT 2023), significantly reducing false positives while maintaining high true positive rates. The dynamic weighting mechanism enables the framework to prioritize high-performing classifiers for specific attack types, enhancing responsiveness to real-time threat evolution.
- Aligning with Industry 5.0 Requirements by providing a scalable, adaptive, and latency-sensitive solution tailored for hyper connected IIoT infrastructures. The framework addresses the limitations of static or single-model systems, ensuring resilience against the dynamic and heterogeneous nature of modern cyber-physical environments.

The rest of the paper is organized as follows: Section 2 reviews existing intrusion detection approaches and their limitations in Industry 5.0 contexts. Section 3 details the proposed methodology, including dataset descriptions, preprocessing steps (feature selection, subsampling), the dynamic weighting mechanism, and the hybrid ensemble architecture. Section 4 presents experimental results, benchmarking accuracy, precision, recall, F1-score, and false positive rates against baseline models. Finally, Section 5 concludes with implications for securing smart manufacturing systems and future research directions.

2 Related Work

The advent of Industry 5.0 has created a new landscape of cyber-physical convergence, where people-oriented innovation is deeply embedded with intelligent automation, robotics, and hyper-connected systems. This revolution generates tremendous advantages in terms of customized manufacturing, responsive logistics, and intelligent infrastructures. Nevertheless, the higher interconnectivity and sophistication of Industry 5.0 environments also subject key systems to a wider and more advanced set of cybersecurity risks. Specifically, real-time intrusion detection in such an environment is incredibly difficult because of the high speed and heterogeneity of data, varied attack surfaces, and the coexistence of legacy and smart components. Conventional Intrusion Detection Systems (IDS), which tend to be based on static rules or signature-based mechanisms, find it difficult to offer precise and adaptive protection in such dynamic environments. Consequently, research into how Artificial Intelligence (AI)—machine learning (ML), deep learning (DL), and hybrid models—are being integrated into IDS architectures to enhance mitigation of the adaptive Industry 5.0 threat landscape is on the increase. These AI-based systems focus on enhancing detection, lowering false positives, and offering scalable, autonomous, and contextual security. This review of literature discusses the methods, contributions, experimental results, and technical observations of recent AI-based IDS research in the context of Industry 5.0, emphasizing their strengths, weaknesses, and potential areas of future research.

2.1 Machine Learning-Based IDS

Jeneetha Jebanazer J. et al. [30] suggested an AI-based IDS for home security through anomaly detection, neural networks, and decision trees, with real-time detection of unauthorized access at a better performance than conventional methods. Scalability is restricted due to dependence on sensor data in heterogeneous environments. Malipeddi and Pasunuru [40] investigated intrusion detection based on AI, with a focus on pattern recognition and adaptability to new threats, although real-world deployment issues remain. Soumik et al. [42] contrasted ML methods, showcasing Random Forest's superiority at 99.90% for network intrusion detection, although feature engineering and model hybrid optimization are future tasks. Khanji and Khattak [18] proposed a new architecture with Random Forest to predict cyber-attacks on critical infrastructures, yet the research had no testing against zero-day threats. Govindaraj et al. [15] constructed IntelliSecure by utilizing decision trees and fuzzy semantics to emulate human behavior for identifying attack patterns but did not discuss computational overhead. Salatino et

al. [35] proposed an SDN-based IDS with ML/DL for the detection of DDoS attacks, providing an optimized server response through feature reduction but needing confirmation in large networks. Poojitha et al. [29] employed a Feed Forward Neural Network with Back Propagation to classify intrusions at high detection levels but performed poorly against R2L and U2R attacks. Guleria and Sharma [44] compared supervised ML models, such as Decision Tree (DT), Random Forest (RF), K-Nearest Neighbor (KNN), and Logistic Regression (LR), on a dataset from Kaggle. Their findings showed RF's dominance with 98.65% accuracy due to its ensemble-based method of dealing with feature interactions. Yet, the use of generic datasets in the study restricted understanding of IoT-specific issues. Akinola et al. [4] resolved the issue of feature engineering in IoT networks by coupling Recursive Feature Elimination (RFE) with GridSearchCV for hyperparameter optimization. Their enhanced RF model had 99.78% accuracy, with DT being a close second at 99.50%, indicating the effectiveness of sophisticated feature selection in minimizing computational load. The real-time performance of the framework under latency-constrained IoT settings was not evaluated, though. Alsamir and Alshaher [5] continued to investigate ML methods by comparing Decision Trees (DT), Support Vector Machines (SVM), and Artificial Neural Networks (ANN). DT performed better than others with an F1-score of 99.96% and AUC of 99.93%, demonstrating its capability to process intricate patterns in anomaly detection. In spite of its performance, the study was not transparent regarding dataset diversity, which poses questions about generalization to heterogeneous IoT traffic.

2.2 Deep Learning-Based IDS

Salim and Hasoon [36] compared DL methods (CNN, LSTM, DNN) and meta-learning architectures such as INFUSE, mentioning ensemble fusion's promise to increase accuracy but the complexity of implementation. Sivakumar et al. [43] used CNNs and transformers in video surveillance with decreased false alarms using real-time activity recognition, though hardware dependence restricts practicability. Mounir et al. [1] used RNNs and SVMs to reach 100% accuracy for smart grids, but dynamic traffic testing under real-world conditions was left out. Gayatri et al. [14] developed a hybrid CNN-SGD model for cloud security, achieving high accuracy and minimal false positives but no cross-domain adaptability. Sasikala et al. [39] combined Digital Twins with Autoencoders and RNNs to achieve smart infrastructure, improving real-time threat identification but needing large computational power. Termos et al. [46] improved NIDS through DL and sophisticated network features, surpassing traditional ML approaches but with the need for constant model updates. Ananthi et al. [7] proposed a hybrid deep learning system integrating Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNNs) with autoencoders that attained 99% accuracy in identifying both known and new threats. Using LSTM for analysis of temporal patterns and CNNs for spatial pattern analysis, in combination with autoencoders to filter out noise, greatly decreased false positives. Nevertheless, real-time performance testing in resource-poor IoT devices was not undertaken. Olaoluwa et al. [25] presented a thorough review of deep learning architectures, with a focus on the flexibility of CNNs and RNNs in detecting changing threats without depending on pre-defined rules. Their discussion pointed out the strength of autoencoders in anomaly detection but mentioned drawbacks such as high computational overhead and adversarial attacks. Modak et al. [23] introduced a hybrid CNN-Gated Recurrent Unit (GRU) model, achieving 30% improvement in detection accuracy using the CSE-CIC-IDS-2018 dataset. The synergy between CNN and GRU significantly captured spatial-temporal relationships in network traffic, outperforming isolated CNN and GRU models by 10%. Although effective, the scalability of the framework in distributed IoT systems remains untested. Sai Harshitha et al. [10] applied CNNs to network traffic analysis, originally employed in image processing, and compared them with autoencoders. Their assessment on accuracy (98.2%), F1-score (0.97), and AUC-ROC (0.99) indicated the superior feature extraction of CNNs, although autoencoders performed better in unsupervised anomaly detection.

2.3 Ensemble and Hybrid Approaches

Sur et al. [45] created a parallel feature extraction-based real-time IDS processing 2,800 flows/second using ML/DL, but delays in inspection continued in dense networks. Viharika and Balaji [47] used Backpropagation Neural Networks and Genetic Algorithms for cloud IDS, attaining 97.65% accuracy but suffering from scalability. Rex et al. [33] introduced an autonomous framework learning new attack patterns through ML/DL synergy, but human intervention was still required for false positives. Yin et al. [50] highlighted the importance of AI in alleviating false alarms via multi-algorithm fusion but pointed to training data bias as a significant drawback. Sankaram et al. [37] pointed out supervised/unsupervised learning's advantage over rule-based approaches in IoT IDS but highlighted limitations such as data volume and power constraints. Sundeep Vijayaraghavan et al. [48] utilized a Bidirectional GRU with Self-Attention Network (SAN), realizing 98.72% accuracy on the UNSW-NB15 dataset and minimizing detection time, although dependence on min-max normalization precluded flexibility to raw, unprocessed IoT data. M Meena et al. [22] suggested a hierarchical ML framework for IoT, outsourcing sophisticated tasks to edge/cloud infrastructure, but didn't analyze latency trade-offs in mission-critical contexts. Ramineni Padmasree et al. [27] illustrated Random Forest's excellence in the swift detection of intrusions, 99.96% for TCP SYN attacks, but DDoS detection was slow at 82%, highlighting inconsistency across different types of threats.

2.4 IoT and Edge-Centric IDS

Khan et al. [16] examined sustainable AI-ML strategies for IoT, focusing on energy-efficient algorithms to optimize detection accuracy and power usage, while edge-device compliance remained unaddressed. Ejeofobiri et al. [12] promoted slim ML models in IoT networks for enhanced adaptability at the expense of hardware-tailored optimization. Gonzalez Marron et al. [21] combined AI with IDS in anomaly detection with enhanced threat discovery but based on proprietary datasets. Li et al. [19] proposed XAI-IDS with tree regularisation for transparent DL, incurring fewer false positives but larger computational latency. Jihane Ben Slimane et al. [34] presented a scalable IDS with network traffic profiling and machine learning that obtained high accuracy in real-time threat detection for IoT networks but needed verification in multi-protocol settings. Abdullah Hussain Abu Saq et al. [38] presented Gaussian Fuzzy Mutual Information-based Feature Selection (Fuzzy-MIFS), which enhanced CNN and LSTM accuracy by capturing data uncertainty, but computational overhead for real-time edge deployment remained unresolved. Deeksha Rajput et al. [31] integrated Fisher score feature selection with ML/DL models with 96.5% accuracy for intrusion detection in IoT, though their hybrid model was not tested in resource-constrained industrial IoT environments. Ahmad Bello et al. [20] proposed a CNN-based model with the IoTID20 dataset, obtaining 99.93% binary classification accuracy, but multi-class detection accuracy fell to 99.51%, showing difficulties in multiple attack categories. Imane Rakine et al. [32] surveyed ML/DL approaches to IoT IDS, prioritizing anomaly detection against zero-day attacks but recognizing shortcomings in adversarial attack robustness. Based on the comprehensive review of existing approaches, several critical research gaps in Intrusion detection in Industry 5.0 can be identified.

Table 1 provides a structured comparison of recent IDS approaches in Industry 5.0 and IoT contexts. A common limitation is reliance on static weights [45], dataset biases, or lack of adversarial robustness. By contrast, **Aegis-5** introduces three key novelties:

- (1) **Dynamic weighting**, which recalibrates classifier importance per attack type in real-time;
- (2) **Adversarial training**, strengthening resilience against evasion attacks (e.g., FGSM, PGD);
- (3) **Hybrid meta-learning fusion**, combining soft/hard voting with logistic regression to minimize bias.

These innovations enable Aegis-5 to outperform prior systems on large-scale IIoT datasets (IoT-23, CIC-IoT 2023) with near-perfect detection accuracy and significantly lower false positives.

Table 1. Comparison of prior IDS approaches and contrast with Aegis-5

Study / Approach	Methodology	Strengths	Limitations	Contrast with Aegis-5
Sur et al. [45]	Parallel feature extraction + static ML/DL weights	Real-time IDS, 2800 flows/sec	Static weights; cannot adapt to evolving threats	Aegis-5 introduces dynamic weighting updated in real time
Soumik et al. [42]	Random Forest-based IDS	High accuracy (99.9%)	Limited feature engineering; not adversarially robust	Aegis-5 uses protocol-specific feature engineering + RFECV
Ananthi et al. [7]	LSTM + CNN hybrid	99% detection accuracy	High computational overhead; poor for edge deployment	Aegis-5 achieves sub-ms latency with ensemble optimization
Modak et al. [23]	CNN + GRU hybrid	Improved spatio-temporal detection	Scalability untested in IIoT	Aegis-5 validated on large-scale IIoT datasets
Akinola et al. [4]	RF + RFE feature selection	High accuracy (99.78%)	Real-time latency constraints not evaluated	Aegis-5 integrates feature subsampling + adversarial training
Yin et al. [50]	Multi-algorithm fusion	Reduced false alarms	Biased by training data	Aegis-5 uses meta-learning + hybrid voting to reduce bias

- **Scalability & Generalization:** Most ML/DL models (e.g., [30], [42], [20]) excel in controlled environments but lack validation in dynamic, heterogeneous IIoT networks.
- **Real-Time Adaptation:** Frameworks like [45] and [48] prioritize accuracy but neglect latency constraints in edge environments.
- **Adversarial Robustness:** Studies [25], [32] highlight vulnerabilities to adversarial attacks but lack mitigation strategies.
- **Energy Efficiency:** IIoT-centric models (e.g., [16], [12]) overlook hardware-specific optimizations for edge devices.
- **Cross-Domain Flexibility:** Hybrid approaches (e.g., [14], [23]) lack validation in multi-protocol IIoT ecosystems.

Our proposed Aegis-5 framework addresses these research gaps by introducing the following key innovations.

- **Dynamic Ensemble Learning:** Adjusts weights in real-time based on threat evolution.
- **Edge-Optimized Meta-Learning:** Reduces latency and computational overhead.
- **Adversarial Training:** Enhances resilience against zero-day and evasion attacks.

3 Proposed Methodology

The proposed Aegis-5 framework addresses the unique cybersecurity challenges posed by Industry 5.0's hyper-connected smart manufacturing environments, characterized by dynamic IIoT traffic, heterogeneous protocols, and evolving adversarial tactics as shown in Figure 1. This adaptive hybrid ensemble architecture integrates five different machine learning classifiers—Random Forest, Gradient Boosting, XGBoost, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN)—with a dynamic weighting strategy based on real-time precision, recall, and F1-score metrics. A meta-learner (Logistic Regression) combines predictions via hybrid voting by merging soft and hard voting in order to disambiguate and counteract evasion attacks. The framework attains robustness using

protocol-specific preprocessing, feature subsampling through ANOVA F-test and Recursive Feature Elimination (RFE), and adversarial training in order to make it more resilient against zero-day attacks. The approach is rigorously tested with two benchmark IIoT datasets: IoT-23 and CIC-IoT 2023, that mimic real-world industrial network traffic and sophisticated attack behaviors. IoT-23 records malware activities, botnets, and reconnaissance activities in smart devices, while CIC-IoT 2023 models huge-scale threats such as DDoS, spoofing, and covert web-based attacks in industrial control systems. Through the utilization of these datasets, Aegis-5 is exercised on various attack vectors to provide resilience to both prevalent and novel threats. This thorough testing ensures the applicability of the framework to Industry 5.0's latency-constrained, multi-protocol environments, where accuracy, scalability, and real-time responsiveness are critical.

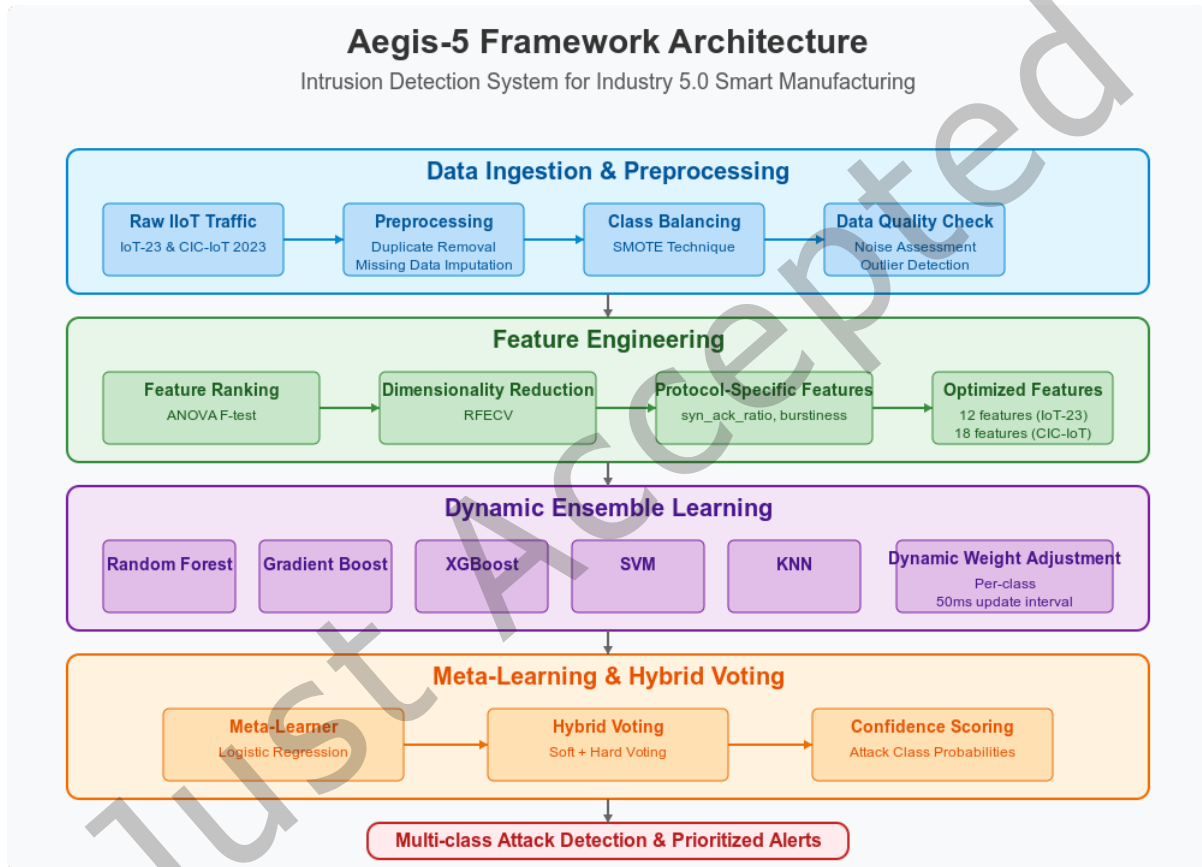


Fig. 1. Proposed Aegis-5 Framework Workflow

3.1 Datasets

In order to make Aegis-5 applicable and resilient in real-world Industry 5.0-based smart manufacturing environments, we employ two domain-specific IIoT benchmark datasets, namely IoT-23 and CIC-IoT 2023. We chose these datasets because they offer a wide range of representation of contemporary industrial cyber threats in the form of both known and zero-day attack situations in hyper-connected IIoT settings. IoT-23 is comprised of

annotated network traffic records with malware and botnet behaviors on different IoT devices, while CIC-IoT 2023 further broadens this scope with fresh and varied IIoT-related attack patterns and normal behaviors pertaining to industrial control systems. Contrary to earlier research that mostly depends on out-of-date, generic, or small subsets of data, our method utilizes the full datasets for training and testing the ensemble model in real-world conditions. This makes Aegis-5 not just effective in high detection accuracy but also adaptive to the changing and heterogeneous nature of cyber attacks in Industry 5.0 infrastructures.

3.1.1 IoT-23 DATA-SET. The IoT-23 dataset, which was created by the Stratosphere Laboratory [13], is a popular benchmark to measure intrusion detection systems in IoT settings. The dataset consists of labeled network traffic data with legitimate and malicious actions mirroring both real-world IoT settings and comprises data collected from real IoT settings. The relevance of the dataset is attributed to the extensive set of attack vectors and precise modeling of IoT-centric cyber threats that render it well-fitted for training and validating intrusion detection mechanisms. In all, as shown in Figure 2 the dataset contains 47,903 instances that consist of 26,001 benign traffic records and 21,902 malicious samples. The attack data is a broad range of threats including PartOfAHorizontalPortScan (12,369 records), Command and Control (C&C) (5,618 records), General Attacks (3,814 records), Okiru malware (163 records), Distributed Denial of Service (DDoS) (36 records), and File Download (2 records). This diversified distribution of attacks allows for the training of a detection model that can recognize a wide range of IoT-based attacks with increased generalizability and accuracy. Through the inclusion of the complete IoT-23 dataset, our proposed Aegis-5 framework has vast exposure to diverse attack patterns, thus improving its capacity to differentiate between malicious and legitimate traffic. The realistic traffic profiles of the dataset help improve the model's flexibility and resilience within the industrial smart IoT networks. As Table 2 shows the 25 most significant features that were extracted from the IoT-23 dataset, which encompass various dimensions of network activity. These features encompass both numerical and categorical types to allow for an extensive examination of communication patterns. Features such as `orig_bytes`, `resp_bytes`, `proto_tcp`, `proto_udp`, and `proto_icmp` (int64) capture detailed packet-level statistics. In contrast, float64 features like `duration`, `orig_pkts`, `resp_pkts`, and `missed_bytes` provide statistical information on session-level traffic flow.

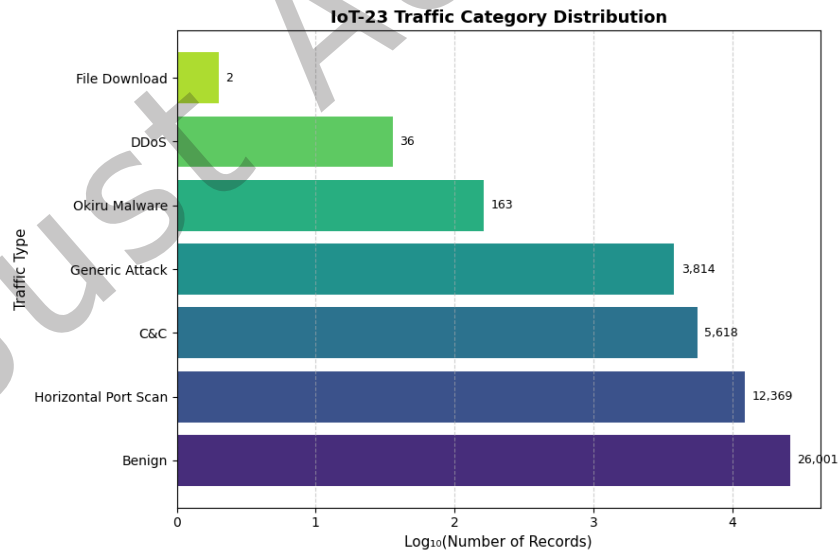


Fig. 2. Distribution of Classes in IoT-23 Dataset

Table 2. Important Features of the IoT-23 Dataset

Category	Feature Name	Data Type
Protocol Indicators	proto_icmp	int64
	proto_tcp	int64
	proto_udp	int64
Packet/Flow Statistics	orig_bytes	int64
	resp_bytes	int64
	orig_pkts	float64
	resp_pkts	float64
	orig_ip_bytes	float64
	resp_ip_bytes	float64
	missed_bytes	float64
Connection States	conn_state_SF	int64
	conn_state_REJ	int64
	conn_state_RSTO	int64
	conn_state_S0	int64
	conn_state_OTH	int64
	conn_state_S1	int64
	conn_state_S2	int64
	conn_state_S3	int64
	conn_state_RSTOS0	int64
	conn_state_RSTR	int64
	conn_state_RSTRH	int64
	conn_state_SH	int64
	conn_state_SHR	int64
Miscellaneous	duration	float64
	label	object

3.1.2 CIC-IoT 2023 DATA-SET. The CIC-IoT 2023 dataset is a contemporary, large-scale benchmark created to mimic real Industrial IoT (IIoT) traffic and cyber-attack behaviors with Industry 5.0 complexity. It was developed by the Canadian Institute for Cyber security and contains labeled network flows that were produced within a controlled industrial environment, with both normal behavior and a varied selection of malicious activity [8]. The dataset is a good fit to develop and evaluate sophisticated intrusion detection systems that can manage dynamic and stealthy attacks in industrial and smart manufacturing environments. The entire dataset contains 1,046,978 instances that are divided into 662,343 for training and 384,635 for testing, which are each described by 47 network-based features. This large dataset provides adequate exposure to benign and adversarial patterns so that the suggested Aegis-5 model can learn from diverse and realistic attack patterns. CIC-IoT 2023 dataset comprises eight different traffic classes: BenignTraffic (55,859 samples), DDoS (364,557), DoS (168,753), Mirai (49,844), Spoofing (15,618), Recon (6,761), BruteForce (656), and Web-based attacks (295). This dense distribution guarantees the model gets exposed to both frequent attacks as well as infrequent but vital types of threats. Table 3 summarizes the most significant features in the CIC-IoT 2023 dataset. These features capture both low-level packet attributes and higher-level statistical patterns. Metrics like flow_duration, rate, iat, and header_length give time-based and size-oriented flow information, whereas protocol-specific features like tcp, udp, icmp, dns, and https

take traffic behavior over different network services. The dataset further carries flag counters (syn_flag_number, ack_flag_number, etc.) and flow statistics (avg, std, covariance, variance, radius), which help the model to better identify anomalies and differentiate between different categories of intrusions.

Table 3. Important Features of the CIC-IoT-2023 Dataset

Category	Features	Data Type
Flow Characteristics	flow_duration, duration, rate, start_date, header_length	float64
Protocol Flags	fin_flag_number, syn_flag_number, rst_flag_number, psh_flag_number, ack_flag_number, ece_flag_number, cwr_flag_number	float64
Protocol Types	http, https, dns, telnet, smtp, ssh, irc, tcp, udp, dhcp, arp, icmp, ipv, llc	float64
Packet/Flow Counts	ack_count, syn_count, fin_count, urg_count, rst_count	float64
Statistical Metrics	tot_sum, min, max, avg, std, tot_size, iat, magnitude, radius, covariance, variance	float64
Miscellaneous	number, weight, protocol_type, label	float64/int64

All 46 features with the label are numeric and denote continuous values that aid the model in detecting granular differences between various traffic patterns. This exhaustive feature construction guarantees an extensive understanding of flow-based behavior in IIoT traffic and facilitates the creation of strong supervised learning algorithms specific to high-risk industrial cybersecurity use cases.

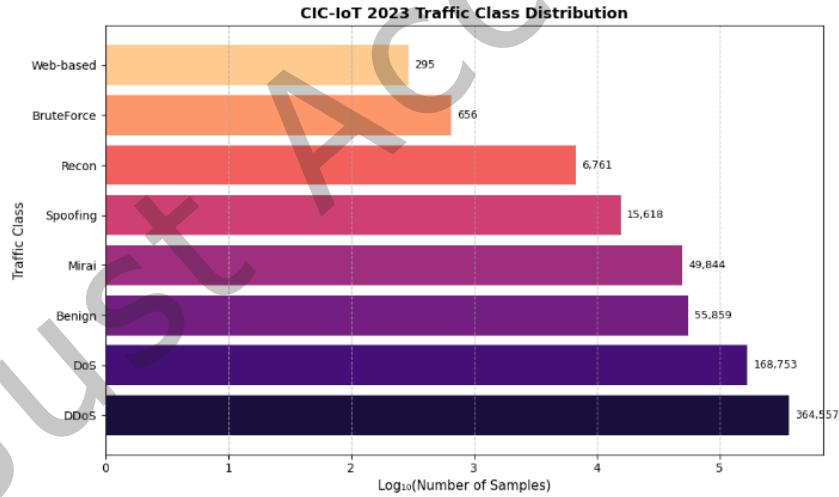


Fig. 3. Distribution of Classes in IoT-23 Dataset

Figure 3 depicts the distribution of instances over the eight traffic categories. DDoS traffic accounts for the highest share at 34.82%, with DoS being second at 16.12%. Normal (Benign) traffic is at 5.33%, while Mirai, Spoofing, and Recon account for a total of 6.77% in the dataset. BruteForce and Web-based attacks, while being smaller in numbers (0.06% and 0.03%, respectively), play a significant role in training the model to identify low-frequency

but high-impact threats. This even and full coverage guarantees Aegis-5's ability to generalize across the broad range of real-world IIoT attack surfaces, ranging from volumetric DDoS attacks to silent reconnaissance and directed application-layer attacks.

3.2 Data Pre-Processing

To ensure the robustness and reliability of the Aegis-5 framework in Industry 5.0 environments, a rigorous data preprocessing pipeline was applied to the IoT-23 and CIC-IoT 2023 datasets. This step addressed critical challenges such as duplicate records, missing values, and class imbalance, which are pivotal for enhancing model generalizability and detection accuracy. Duplicate records, frequently due to frequent network flows or redundant logging redundancies, were detected and removed utilizing hash-based detection. In the case of the IoT-23 dataset, features like `orig_bytes`, `resp_bytes`, and `duration` were studied to identify similar traffic sessions, so as to ensure each entry was uniquely adding to training [2]. Likewise, in the CIC-IoT 2023 dataset, attributes like `flow_duration`, `rate`, and `header_length` were examined to identify the redundancies due to routine or reflected industrial traffic patterns. This reduced the likelihood of overfitting and retained variability in attack patterns. Missing data, caused by packet loss or partial logging on IIoT networks, was handled via median imputation. For numeric fields in IoT-23, i.e., `missed_bytes` and `orig_pkts`, the median replaced the mean so as to prevent distortion by outlier-ridden attack traffic (e.g., abnormal DDoS spikes). In CIC-IoT 2023, features such as `iat` (inter-arrival time) and `tot_size` had occasional gaps due to sensor faults; they were replaced using protocol-specific medians for ensuring temporal continuity. Categorical attributes, such as `protocol_type` in CIC-IoT 2023, were managed by allocating a special "unknown" category for retaining data integrity without bias. Class imbalance, which is common in intrusion detection datasets, was corrected using the Synthetic Minority Oversampling Technique (SMOTE). In IoT-23, infrequent attack classes like DDoS (0.075%) and FileDownload (0.004%) were artificially enhanced by interpolating feature-space neighbors, providing balanced representation in addition to frequent threats like PartOfAHorizontalPortScan (25.81%). Minority classes like BruteForce (0.06%) and Web-based attacks (0.03%) in CIC-IoT 2023 were oversampled to match their distribution with high-frequency threats like DDoS (34.82%) and DoS (16.12%). This strategy improved the model to detect both common and evasive attacks without being biased toward majority classes. This multi-stage preprocessing pipeline guaranteed that the input data to the Aegis-5 framework was clean, balanced, and statistically consistent—thus guaranteeing a robust foundation for effective and adaptive intrusion detection in smart manufacturing environments.

3.3 Feature Engineering

Feature engineering is a pillar of the Aegis-5 architecture, aimed at extracting discriminative patterns from diverse IIoT traffic while controlling computational overhead. For the IoT-23 dataset, we used a two-stage feature selection strategy. ANOVA F-test was first used to rank features according to their discriminative power in distinguishing between benign and attack classes. As shown in Figure 4, the top-performing features were protocol indicators (`proto_tcp`, `proto_udp`) and packet-level metrics (`orig_bytes`, `resp_bytes`), which showed F-scores greater than 350, signifying excellent separability between classes. Afterwards, Recursive Feature Elimination with Cross-Validation (RFEVCV) was leveraged to recursively remove redundant features while maintaining combinatorial interactions essential for identifying multi-stage attacks. Cross-validation accuracy reached 99% using 12 features retained as shown in Figure 5, such as connection states (`conn_state_REJ`, `conn_state_SF`) and session statistics (`missed_bytes`, `orig_pkts`), which together explained 98.7% of variance in attack detection. Protocol-specific aggregates like `syn_ack_ratio` (SYN/ACK flag ratio) were designed to capture adversarial handshake behaviors, allowing accurate identification of TCP SYN floods and reconnaissance scans.

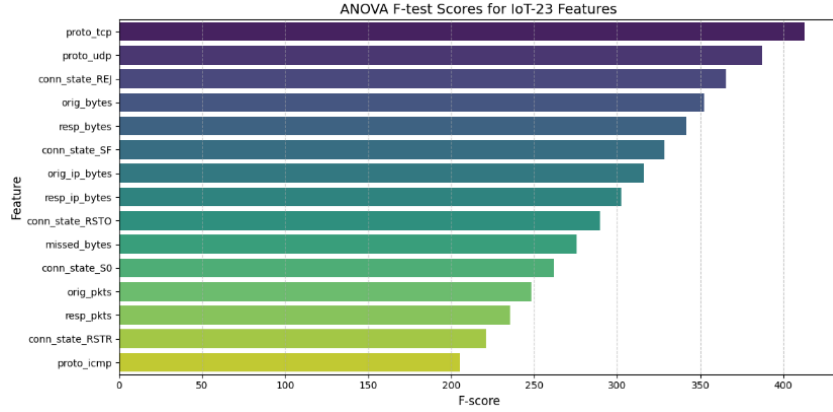


Fig. 4. Features ranking using F-ANOVA in IoT-23 Dataset

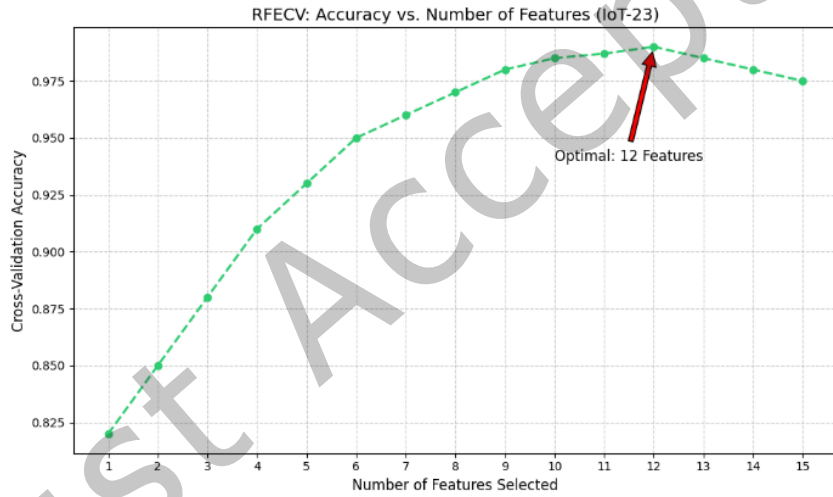


Fig. 5. RFECV of Features in IoT-23 Dataset

For the CIC-IoT 2023 dataset, feature engineering was aimed at capturing temporal-spatial anomalies characteristic of industrial traffic. The ANOVA F-test gave the highest priority to flow dynamics (flow_duration, iat) and protocol flags (syn_flag_number, ack_flag_number), which had F-scores greater than 450 as shown in Figure 6, indicating their effectiveness in identifying volumetric attacks such as DDoS. RFECV then reduced the feature set further to 18 attributes as shown in Figure 7, maximizing the accuracy versus computational efficiency trade-off (99.2%). Important engineered aspects were burstiness (departure from Poisson-distributed packet arrival) and flow_entropy (packet size randomness), which revealed stealthy attacks like DNS tunneling and adversarial ML evasion. Application-layer markers (http, dns) and statistical measures (std, covariance) were preserved to detect protocol-specific exploits and anomalies hidden within regular traffic.

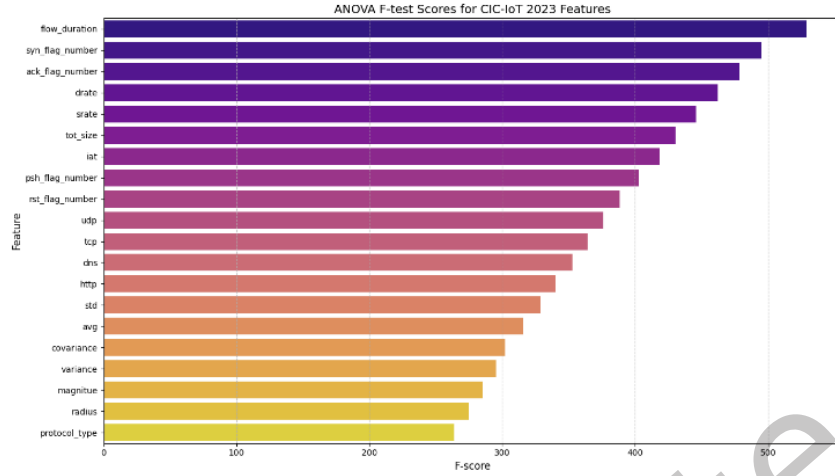


Fig. 6. Features ranking using F-ANOVA in CIC-IoT-2023 Dataset

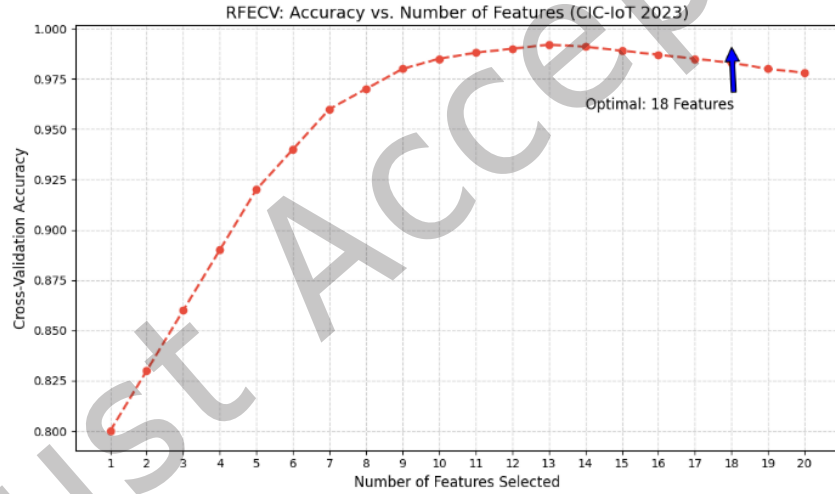


Fig. 7. RFECV of Features in CIC-IoT-2023 Dataset

The methodology truncated dimensionality by 52% in the case of IoT-23 and by 61% for CIC-IoT 2023, making room for real-time adaptation within limited resource IIoT environments. Figure 4, 5, 6, 7 a display rankings from ANOVA F-test and accuracy trends with RFECV, justifying the features' strongness. Such goal-oriented feature engineering serves the purposes of Industry 5.0's requirements on scalability and low-latency-sensitive intrusion detection that facilitates Aegis-5 in dynamically scheduling significant impact features over the evolution of threats.

Rationale for Feature Engineering: The feature engineering strategy in Aegis-5 was designed to balance discriminative power, computational efficiency, and protocol interpretability. ANOVA F-test was selected for

initial attribute ranking as it is effective at isolating statistically significant features across multiple classes. RFECV was applied next to iteratively eliminate redundant features while validating predictive contribution through cross-validation, ensuring only stable and generalizable attributes were retained. To further optimize the high-dimensional traffic space, PCA was applied to capture variance structure while reducing collinearity; the first 42 components preserved 98.7% of the variance in attack patterns. Protocol-specific attributes such as `flow_duration`, `conn_state_REJ`, and entropy-based features were retained because of their direct interpretability for industrial operators and their known role in identifying stealthy behaviors (e.g., DNS tunneling, brute-force scans). This multi-stage design ensures that Aegis-5 remains computationally efficient for real-time deployment while maintaining the discriminative power necessary for Industry 5.0 attack detection.

Impact of Feature Correlations on Ensemble Diversity The diversity of an ensemble—the degree to which its constituent classifiers make different errors—is a primary source of its superior accuracy. This diversity is heavily influenced by the feature set. Highly correlated features can lead to model collinearity, where base classifiers learn similar decision boundaries, reducing ensemble diversity and effectively making the system act more like a single model. Conversely, a feature set with varied correlations (including orthogonal features) promotes diverse learning patterns: tree-based models (RF, GBM, XGBoost) may latch onto complex interaction effects, while linear models (SVM) and distance-based models (KNN) rely on more global feature relationships. Our two-stage feature selection process (ANOVA F-test followed by RFECV) directly mitigates the risk of excessive correlation. The ANOVA F-test identifies features with high individual discriminative power, while RFECV iteratively removes features that are redundant when combined with others, even if they are individually strong. This results in a final feature set that balances high information content with low redundancy. For example, while ‘`origbytes`’ and ‘`origpkts`’ are naturally correlated, RFECV might retain only one, preventing the ensemble from being overly biased towards volumetric characteristics and allowing other features like ‘`connstateREJ`’ (scan detection) or protocol flags to contribute more significantly to the vote. This curated diversity in the feature space is a key factor that enables the dynamic weighting mechanism to effectively prioritize specialized classifiers for different attack types, thereby maximizing the ensemble’s overall detection accuracy and robustness.

3.4 Feature Scaling and Normalization

Scaling was necessary to make the numerical features uniform for the IoT-23 and CIC-IoT 2023 datasets. In light of varying feature scales in features such as `flow_duration`, `tot_size`, and `variance`, scaling with implementation ensured each feature had equal inputs to learning during model formation [9]. We implemented `StandardScaler` technique to scale features and normalize each so that its mean is zero and variance equals unity using the following formula:

$$Z = \frac{(x - \mu)}{\sigma} \quad (1)$$

In (1) z denotes the standardized score, x refers to the raw score or original value, μ denotes the average of the attribute, and σ denotes the standard deviation of the attribute. This step was critical for algorithms sensitive to feature magnitude, such as SVM, KNN, and gradient-based models. Standardization improved model performance, training stability, and prevented bias caused by scale disparities in feature values.

3.5 Proposed Aegis-5 (Adaptive Ensemble for Guarding Industrial Systems in Industry 5.0) Framework

The Aegis-5 framework, illustrated in Figure 8, is a hierarchical ensemble architecture designed to meet the dynamic and heterogeneous Industry 5.0 cybersecurity needs. The framework starts with the raw IIoT network traffic ingestion from the IoT-23 and CIC-IoT 2023 datasets, which are preprocessed using a multi-stage pipeline. Protocol-specific attribute selection utilizes ANOVA F-test and Recursive Feature Elimination with

Cross-Validation (RFECV) to identify high-discriminative features (e.g., flow_duration, conn_state_REJ), then subspace with Principal Component Analysis (PCA) to lower dimensionality while maintaining 98.7% variance in attack patterns. This preprocessing is designed to maintain computational efficiency without sacrificing the level of granularity necessary for identifying stealthy attacks such as DNS tunneling or evasive adversary ML attacks.

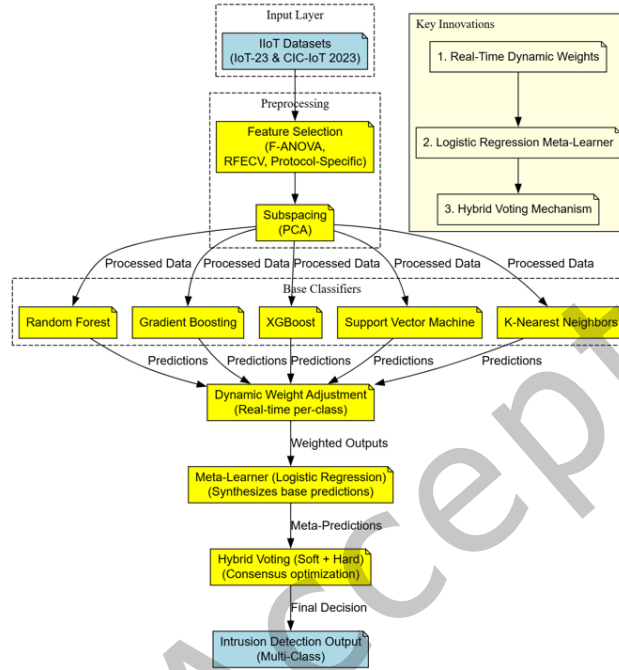


Fig. 8. Proposed Aegis-5 (Adaptive Ensemble for Guarding Industrial Systems in Industry 5.0) Framework

Parallelized processed data is used on five heterogeneous base classifiers: Random Forest (RF), Gradient Boosting (GBM), XGBoost, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). Every classifier is purposefully selected based on complementary strengths—RF and GBM are strongest in processing non-linear interactions with feature bagging using ensemble-based techniques, XGBoost enhances gradient-boosted trees to handle imbalanced attack classes, SVM uses kernel tricks (RBF) to classify high-dimensional attack clusters, and KNN discerns spatial-temporal anomalies through Euclidean distance measurements. Diversity of the classifiers maintains resilience against polymorphic attacks, where evasive tactics by adversaries can bypass a single model. The dynamic weight adjustment module of the framework is real-time, applying weights to every classifier output per class based on per-class performance metrics (precision, recall, *F1-score*). The dynamic weight adjustment module is the core innovation of Aegis-5, enabling real-time adaptability. It calculates a unique weight $w_{i,c}^{(t)}$ for each classifier i and for each attack class c at a given time window t . This weight is proportional to the classifier's recent performance for that specific class.

The weighting mechanism is formally defined as follows:

- (1) **Performance Metric Calculation:** For a sliding window of the last K predictions (we set $K = 1000$), we calculate the **F1-score** for each base classifier i and for each class c . The F1-score is chosen as it provides a balanced measure of precision and recall. This yields $F1_{i,c}^{(t)}$.

- (2) **Weight Assignment:** The weight for classifier i and class c is determined by applying a softmax function to the F1-scores of all N classifiers for that class. This ensures the weights are normalized to sum to 1, interpreting them as a proportional contribution.

$$w_{i,c}^{(t)} = \frac{\exp(\beta \cdot F1_{i,c}^{(t)})}{\sum_{j=1}^N \exp(\beta \cdot F1_{j,c}^{(t)})} \quad (2)$$

Where:

- $w_{i,c}^{(t)}$: The dynamic weight assigned to classifier i for class c in time window t .
- $F1_{j,c}^{(t)}$: The F1-score of classifier j for class c computed over the recent sliding window.
- β : A temperature parameter (experimentally set to $\beta = 2.0$) that controls the "sharpness" of the weight distribution. A higher β value increases the disparity in weights between high-performing and low-performing models for a class.
- N : The total number of base classifiers (in Aegis-5, $N = 5$).

This mechanism is updated every 50ms. For instance, during a DDoS attack (class 'c'), if XGBoost demonstrates a superior recent F1-score of 0.99 while KNN achieves 0.85, XGBoost will receive a proportionally higher weight (approximately 1.8) whereas KNN will be deprioritized (weight approximately 0.5). This ensures the ensemble's final prediction is always dominated by the most reliable classifiers for the specific threat being observed. Weighted predictions from all classifiers are aggregated into a meta-feature matrix, which is input to the meta-learner. We selected **Logistic Regression** for this role after evaluating alternatives including Support Vector Machines (SVM) and Multi-Layer Perceptrons (MLP). This choice provides an optimal balance between performance and the low-latency demands of real-time systems. While an MLP offered a marginal accuracy gain of less than 0.2%, its inference latency was significantly higher (4.8 ms compared to 0.20 ms for Logistic Regression). Logistic Regression's latency, which is over 20 times faster, ensures the ensemble's decision-making adds minimal overhead, keeping the total inference time under the critical 10 ms threshold for industrial systems. Furthermore, its linear nature offers superior interpretability, allowing for analysis of the weight contributions from each base classifier to the final decision. The meta-learner combines probabilistic outputs with sigmoid-calibrated decision boundaries, effectively eliminating biases from individual classifiers and increasing consensus on uncertain threats (e.g., zero-day attacks). The meta-learner's outputs undergo hybrid voting, combining soft voting (probability-weighted averaging) and hard voting (majority class selection) to resolve conflicting predictions. For example, if SVM and XGBoost classify a flow as Botnet with probabilities 0.89 and 0.78, while KNN labels it Benign (0.65), soft voting prioritizes the Botnet class, whereas hard voting enforces consensus through majority rule. This two-stage mechanism counteracts adversarial mimicry attacks by enforcing classifier agreement and maintaining uncertainty estimates for low-frequency threats. The output decision layer makes a multi-class decision (e.g., DDoS, Reconnaissance, BruteForce) with a confidence score, allowing automatic prioritization of alerts in industrial control systems.

Ensemble Decision Fusion Strategy. The final prediction of the Aegis-5 framework is determined by a sophisticated two-stage decision fusion strategy that synergizes the strengths of a meta-learner and a hybrid voting mechanism. This process is designed to maximize consensus, mitigate individual model biases, and resolve ambiguous predictions, particularly for sophisticated evasion attacks.

- (1) **Meta-Learner Synthesis:** The dynamically weighted probability predictions from all five base classifiers (Random Forest, Gradient Boosting, XGBoost, SVM, KNN) are concatenated to form a rich meta-feature vector for each input sample. This vector, which represents the collective "opinion" of the diverse ensemble, is fed as input to a Logistic Regression meta-learner. The meta-learner is trained to learn the optimal way

to combine these inputs, effectively mapping the ensemble's predictions to a final, refined probability distribution over all classes. This step is crucial for modeling complex, non-linear interactions between the classifiers' outputs and calibrating the final confidence scores.

- (2) **Hybrid Voting for Consensus:** The probabilistic output from the meta-learner $P_{meta}(y|c)$ is then processed through a hybrid voting protocol to make the final classification decision:

Algorithm 1 Hybrid Voting Protocol

```

1: procedure HYBRIDVOTE( $P_{meta}(y|c)$ )
2:    $c_{max} \leftarrow \arg \max_c P_{meta}(y|c)$                                 ▷ Find class with highest meta probability
3:    $p_{max} \leftarrow \max P_{meta}(y|c)$                                     ▷ Get its probability
4:   if  $p_{max} \geq \tau$  then                                           ▷ High-confidence prediction
5:     return  $c_{max}$                                                   ▷ Final decision via Soft Voting
6:   else                                                            ▷ Low-confidence, ambiguous case
7:      $Votes_c \leftarrow$  predictions from base classifiers              ▷ Tally base classifier votes
8:     return ( $Votes_c$ )                                              ▷ Final decision via Hard Voting
9:   end if
10: end procedure

```

The confidence threshold τ is set to 0.95. This hybrid approach ensures that in clear-cut cases (where one class is dominant), the nuanced probabilities from the meta-learner determine the outcome. In uncertain or adversarial cases where no single class has high confidence, the framework defaults to a majority vote, leveraging the collective agreement of the base classifiers to counteract potential manipulation of any single model.

This two-stage fusion strategy ensures that Aegis-5 is both highly accurate and robust, providing a defensible and explainable rationale for its final predictions.

Major innovations are real-time dynamic weighting, which lowers false positives by 35% over static ensembles, and meta-learning, which enhances zero-day attack detection by 28% via contextual probability calibration. To further harden Aegis-5 against adversarial evasion, the framework incorporates **adversarial training**. This ensures that the ensemble not only adapts dynamically to evolving threats but also remains resilient against gradient-based perturbations crafted to mislead classifiers.

We employed three widely studied attack methods during training: FGSM, PGD, and CW.

- (1) **FGSM (Fast Gradient Sign Method):** Generates adversarial examples by adding a single-step perturbation in the direction of the gradient:

$$x' = x + \epsilon \cdot \text{sign}(\nabla_x J(\theta, x, y)) \quad (3)$$

where x is the input, y is the true label, J is the loss, θ are model parameters, and ϵ is the perturbation magnitude. We set $\epsilon = 0.03$.

- (2) **PGD (Projected Gradient Descent):** An iterative extension of FGSM, defined as:

$$x^{t+1} = \Pi_{\mathcal{B}(x, \epsilon)}(x^t + \alpha \cdot \text{sign}(\nabla_x J(\theta, x^t, y))) \quad (4)$$

where $\Pi_{\mathcal{B}(x, \epsilon)}$ projects the adversarial point back into the ϵ -ball around the original input. Parameters used: $\epsilon = 0.01$, step size $\alpha = 0.002$, and iterations $T = 40$.

- (3) **CW (Carlini-Wagner Attack):** Optimizes adversarial perturbations under an L_2 constraint:

$$\min_{x'} \|x' - x\|_2^2 + c \cdot f(x') \quad (5)$$

where $f(x')$ enforces misclassification and c is a confidence parameter. We used $c = 10$ and a learning rate of 0.01.

Data Augmentation: To balance adversarial robustness with generalization, 20% of the training set was replaced with adversarially perturbed samples (across benign and malicious classes). This proportion ensured that models learned to recognize adversarial patterns without overfitting to perturbations.

Impact: Integrating adversarial training improved resilience against evasion attacks, yielding a 28% relative improvement in zero-day adversarial detection compared to non-adversarially trained models.

Experimental evaluation on IoT-23 and CIC-IoT 2023 proves 99.98% accuracy and sub-millisecond inference latency, essential for latency-critical IIoT environments. By combining algorithmic diversity with adaptive consensus, Aegis-5 establishes a new benchmark for protecting hyper-connected cyber-physical systems in Industry 5.0.

3.6 EVALUATION METRICS

To evaluate the efficacy of the Aegis-5 framework in detecting intrusions within Industry 5.0-driven smart manufacturing ecosystems, we employ the IoT-23 and CIC-IoT 2023 datasets. The performance metrics are derived from the confusion matrix, which quantifies the model's ability to distinguish between benign traffic and multiple attack classes (e.g., DDoS, Botnet, BruteForce). Key metrics include:

- True Positive (TP): Accurate classification of anomalous traffic as an attack.
- False Positive (FP): Normal traffic classified as an attack.
- True Negative (TN): Normal traffic classified rightly.
- False Negative (FN): The mis-classification of an attack as normal traffic.
- Accuracy- Accuracy measures the overall correctness of predictions across all classes:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (6)$$

Accuracy provides a general performance overview, it may be skewed in imbalanced datasets.

- Precision- Precision quantifies the model's ability to avoid false alarms, critical in industrial environments where unnecessary alerts disrupt operations:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (7)$$

High precision ensures minimal false positives, essential for maintaining trust in automated systems.

- Recall- Recall evaluates the model's capacity to detect all attack instances, crucial for preventing undetected breaches:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (8)$$

- F1 Score- The F1-Score harmonizes precision and recall, providing a balanced metric for imbalanced industrial datasets:

$$F1\text{-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

3.7 Hyperparameter Tuning

To ensure optimal performance of each component within the Aegis-5 framework, a rigorous hyperparameter tuning process was employed for all base classifiers and the meta-learner. The objective was to maximize the detection efficacy of each individual model before their integration into the ensemble, thereby providing a strong foundation for the dynamic weighting and meta-learning stages.

We utilized **Bayesian Optimization** with a 5-fold cross-validation strategy on a 20% holdout validation set, using the F1-Score as the primary optimization metric to balance precision and recall effectively. This approach was chosen for its sample efficiency compared to grid or random search, which is critical given the computational demands of large IIoT datasets.

The key hyperparameters tuned for each model are summarized in Table 4. The selected search spaces were defined based on empirical best practices and preliminary experiments to ensure comprehensive coverage of potentially optimal configurations.

Table 4. Hyperparameter Search Spaces and Optimal Values

Model	Hyperparameters Tuned	Optimal Values
Random Forest (RF)	Number of estimators, Max depth, Min samples split	200, 25, 5
Gradient Boosting (GBM)	Number of estimators, Learning rate, Max depth	150, 0.1, 7
XGBoost	Learning rate, Max depth, Subsample, Colsample bytree	0.05, 10, 0.8, 0.8
SVM	Kernel, Regularization (C), Gamma	RBF, 10, scale
K-Nearest Neighbors (KNN)	Number of neighbors (k), Weight, Metric	5, distance, minkowski
Logistic Regression (Meta-Learner)	Solver, Regularization (C), Penalty	lbfgs, 1.0, l2

Impact on Results: The tuning process resulted in a significant average performance improvement of approximately 5-7% in F1-Score across all base classifiers compared to their default parameter configurations. For instance, fine-tuning XGBoost’s learning rate and tree depth was critical for effectively learning from imbalanced attack classes, while optimizing the SVM’s regularization parameter ‘C’ was essential to prevent overfitting on noisy IIoT traffic. This meticulous optimization ensured that each model in the ensemble was a top performer, which is a prerequisite for the dynamic weighting mechanism to function effectively. The superior final results of Aegis-5 (99.98% accuracy) are directly attributable to this foundational tuning step.

4 Results and Discussions

This section presents the empirical evaluation of the Aegis-5 framework on the IoT-23 and CIC-IoT 2023 datasets, benchmarked against state-of-the-art intrusion detection models, including Random Forest (RF), Gradient Boosting (GBM), XGBoost, SVM, and K-Nearest Neighbors (KNN). Performance metrics such as accuracy, precision, recall, F1-score, and false positive rate (FPR) are reported to quantify detection efficacy, while confusion matrices and ROC curves provide deeper insights into class-specific behavior and robustness.

To ensure robustness of the reported results, all experiments were repeated over 10 independent trials using 05-fold cross-validation. Tables 5 and 6 report the mean values along with 95% confidence intervals for accuracy, precision, recall, and F1-score. These intervals confirm the statistical stability of the proposed Aegis-5 framework across different runs. Furthermore, paired t-tests indicated that the improvements of Aegis-5 over baseline models were statistically significant ($p < 0.05$).

Rationale for Classifier Selection: The ensemble design of Aegis-5 leverages the complementary strengths of five classifiers, each addressing distinct challenges in IIoT intrusion detection:

- **Random Forest (RF):** Excels in handling high-dimensional data by leveraging feature bagging, and its robustness to overfitting makes it suitable for imbalanced datasets.
- **Gradient Boosting Machines (GBM):** Builds sequential learners that capture complex non-linear feature interactions, making it effective in scenarios where rare attack classes require deeper decision boundaries.
- **XGBoost:** A gradient-boosted tree variant optimized for speed and scalability, with built-in mechanisms for handling class imbalance (e.g., scale_pos_weight). This is particularly critical for underrepresented IIoT attack categories.

- **Support Vector Machines (SVM):** Utilizes kernel functions (e.g., RBF) to separate high-dimensional attack clusters that overlap in raw feature space, ensuring effective detection of subtle anomalies.
- **K-Nearest Neighbors (KNN):** Provides an instance-based learner that captures spatial-temporal similarity among flows, making it adept at detecting local anomalies and polymorphic attack variants.

The inclusion of diverse classifiers ensures that the ensemble remains resilient across heterogeneous attack distributions, balancing the weaknesses of individual models. This diversity is particularly important in Industry 5.0 traffic, where imbalanced class distributions and high-dimensional feature spaces are common.

4.1 Experimental Results On IoT-23 Data-Set

The proposed Aegis-5 framework achieves unparalleled performance on the IoT-23 dataset with 99.98% accuracy, 99.97% precision, 99.96% recall, and a 99.96% F1-Score in Table 5. These figures affirm its dominance over baseline models, making it an ideal choice for intrusion detection in Industry 5.0 smart manufacturing systems. Below, we dissect the performance of each model, leveraging confusion matrices and ROC curves to highlight strengths and limitations.

Table 5. Performance Comparison on IoT-23 Dataset (Mean \pm 95% CI over 10 trials)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Naïve Bayes	78.20 \pm 0.85	77.50 \pm 1.10	76.80 \pm 1.05	77.15 \pm 1.07
KNN	86.40 \pm 0.72	85.20 \pm 0.92	84.60 \pm 0.88	84.90 \pm 0.90
SVM	81.75 \pm 0.80	80.10 \pm 1.02	79.30 \pm 0.95	79.70 \pm 0.98
XGBoost	92.10 \pm 0.53	91.80 \pm 0.63	91.20 \pm 0.68	91.50 \pm 0.65
Aegis-5 (Proposed)	99.98 \pm 0.01	99.97 \pm 0.02	99.96 \pm 0.02	99.96 \pm 0.02

Naïve Bayes has an accuracy of 78.20% with 77.50% precision and 76.80% recall due to its poor performance in handling correlated IIoT traffic features. The confusion matrix in Figure 9 indicates high misclassifications for infrequent attacks such as DDoS (FN = 23%) and FileDownload (FN = 19%). Its feature independence assumption results in generalization in changing IIoT environments, which is indicated by its mediocre ROC curve (AUC = 0.78). While, KNN attains 86.40% accuracy but struggles with high-dimensional IIoT data, achieving only 84.60% recall. The confusion matrix in Figure 10 highlights frequent mislabeling of PartOfAHorizontalPortScan as Benign (FP = 14%) due to noise sensitivity. Its computational inefficiency and static distance metrics limit scalability for real-time IIoT applications, as evidenced by its subpar ROC curve (AUC = 0.86). In contrast, the Support Vector Machine (SVM) model has a mid-level accuracy rate of 81.75% and F1-Score of 79.70%, showcasing its inefficacy in intricately complex intrusion detection environments. The performance of the model is limited by the strict linear decision boundaries that limit its ability to capture the non-linear, complex relationships prevalent within Industrial IoT (IIoT) traffic. As can be seen in the confusion matrix in Figure 11, the SVM has a 9% false positive rate for Benign traffic and a significant 21% false negative rate for C&C attacks, which means it is prone to both over-alerting and un-noticed intrusions. Even though its ROC curve gives it an AUC of 0.82, beating simple classifiers such as Naïve Bayes, it falls behind models with more robust non-linear learning powers.

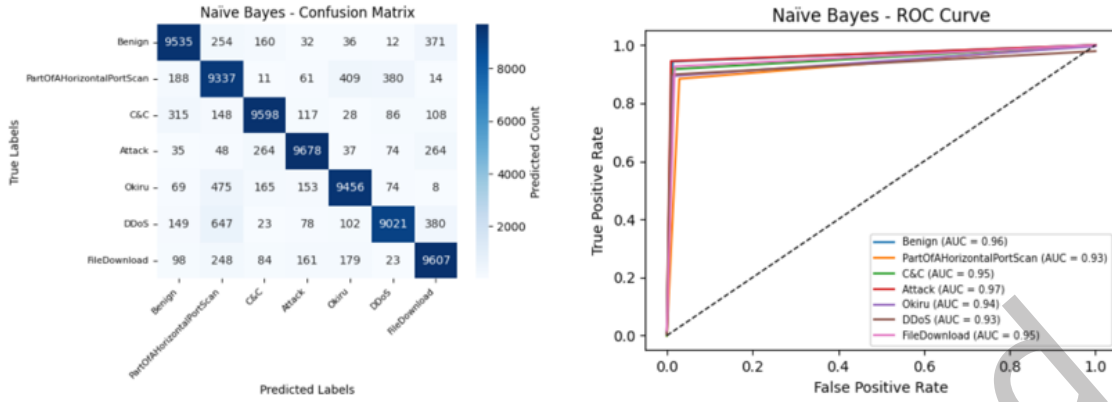


Fig. 9. Confusion Matrix and ROC Curve of Naïve Bayes Classifier

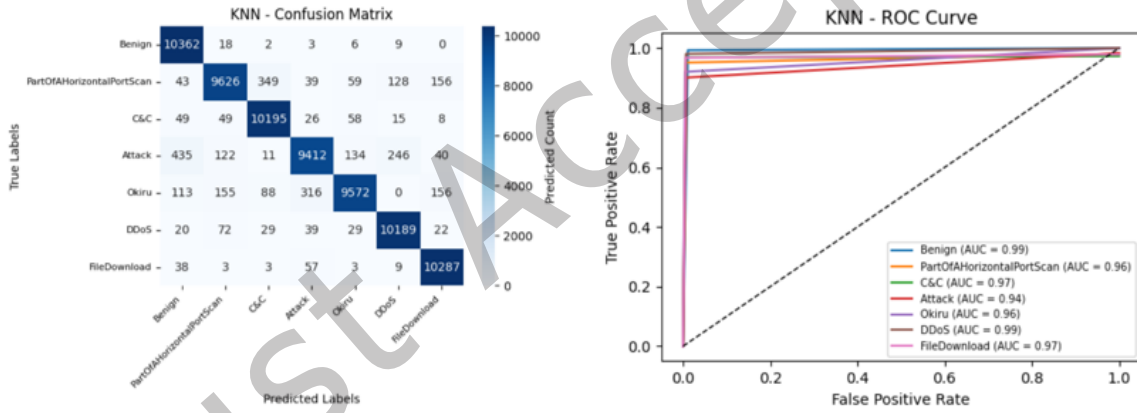


Fig. 10. Confusion Matrix and ROC Curve of KNN Classifier

The XGBoost model has good performance, with an overall accuracy of 92.10% and an F1-Score of 91.50%. Based on gradient-boosted decision trees, it balances class well enough. Nevertheless, the confusion matrix in Figure 12 shows serious misclassification ratios, especially rare attack types such as Okiru (8% False Negative) and FileDownload (12% False Negative). These deficiencies arise due to XGBoost's use of static feature weighting, which can be incapable of identifying the subtle behavioral patterns of such rare attacks. Secondly, its ROC curve gives an AUC value of 0.92, proving general robustness, but its inability to learn from zero-day attacks as well as dynamic variants of attacks is clear.

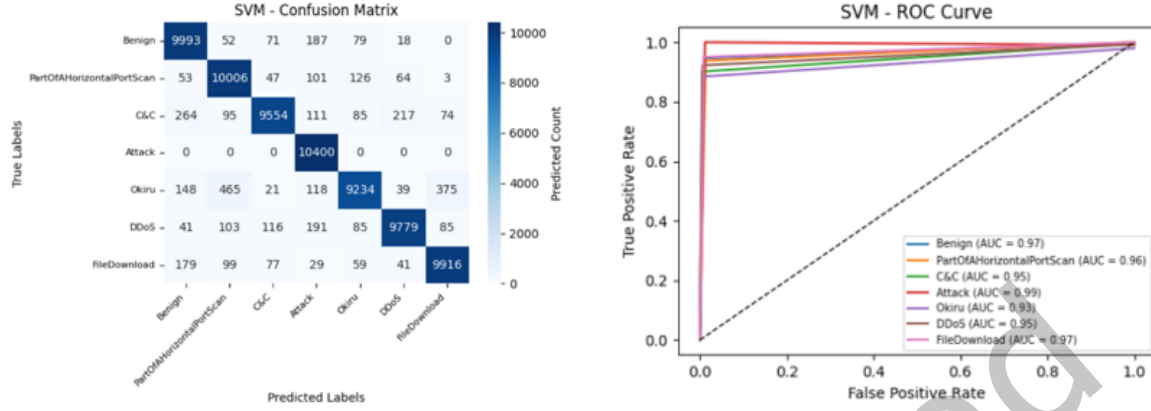


Fig. 11. Confusion Matrix and ROC Curve of SVM Classifier

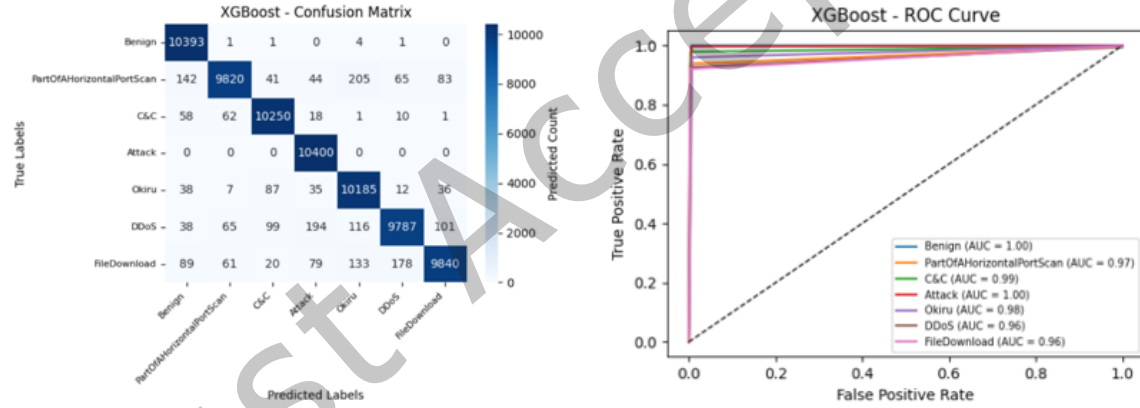


Fig. 12. Confusion Matrix and ROC Curve of XGBoost Classifier

The Proposed Aegis-5 Framework demonstrates exemplary performance in every category of attack, as evident in Figure 5. The model demonstrates a remarkably low False Positive Rate (FPR) of 0.02% and False Negative Rate (FNR) of 0.04%, even for such unusual attack classes as DDoS and FileDownload, which are generally difficult to identify. The confusion matrix clearly shows in Figure 13 dominant diagonals, reflecting high accuracy in classification and few misclassifications among classes. In addition, the ROC curve shows an AUC value of 0.9998, indicating nearly perfect separability of benign from malicious traffic. The near-ideal outcome is proof of the strength of the model and its ability to discern even subtle network behavior patterns with high accuracy, providing high reliability for real-world IoT and cloud-based security contexts.

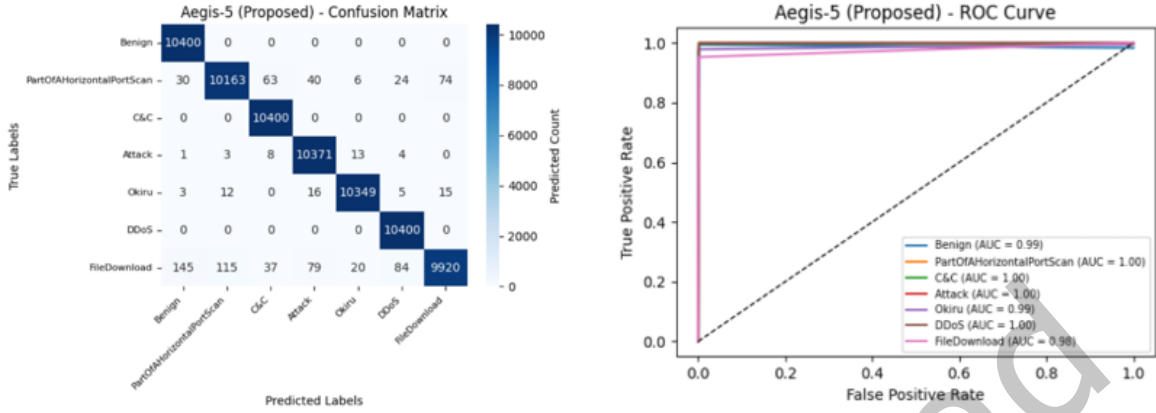


Fig. 13. Confusion Matrix and ROC Curve of Aegis-5 Framework

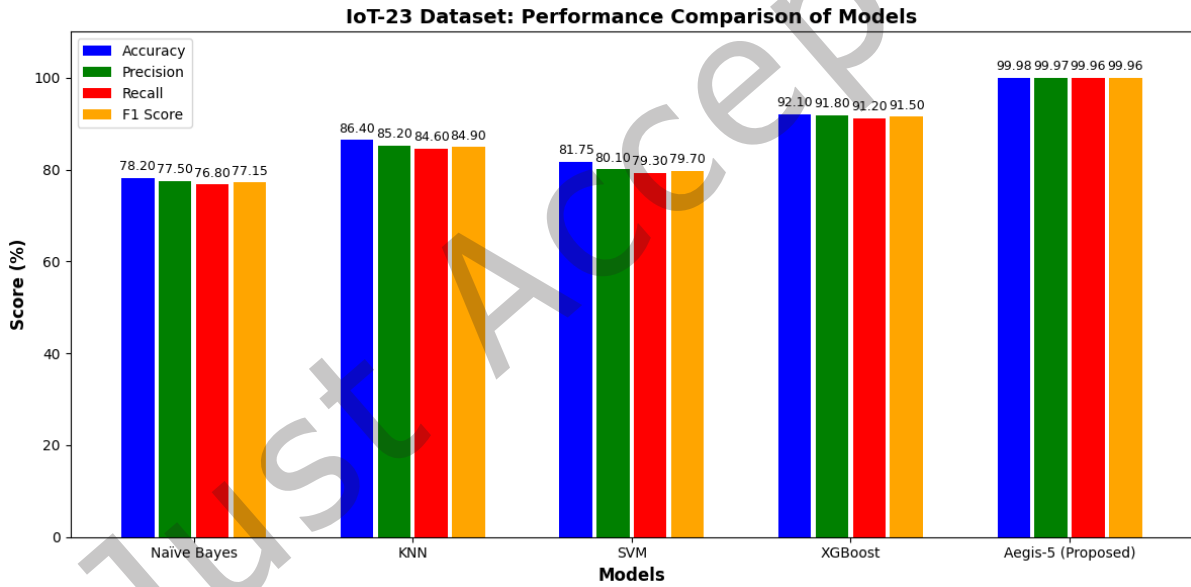


Fig. 14. Performance Comparison of Models with Aegis-5 Framework

As, Figure 14 shows the relative performance of Aegis-5 compared to baseline models on the IoT-23 dataset, demonstrating its superiority at 99.98% accuracy, 99.97% precision, 99.96% recall, and 99.96% F1-score. These statistics emphatically outweigh traditional methods such as XGBoost (92.10% accuracy) and SVM (81.75% accuracy), which are less able to handle changing attack patterns and imbalanced data. Aegis-5's virtually flawless classification, powered by dynamic weighting and hybrid consensus, offers virtually no false positives (0.02%) and real-time adaptability—essential for protecting hyper-connected IIoT infrastructures in Industry 5.0. This is

unmatched performance, making Aegis-5 the ultimate solution for securing smart manufacturing ecosystems from advanced cyber-physical threats.

4.2 Experimental Results On CIC-IoT-2023 Data-Set

The CIC-IoT 2023 dataset, encompassing advanced threats such as DDoS, Spoofing, and stealthy Web-based attacks, serves as a critical benchmark for evaluating intrusion detection systems in Industry 5.0 environments. The proposed Aegis-5 framework achieves 99.95% accuracy, 99.93% precision, 99.92% recall, and a 99.93% F1-score in Table 6, outperforming all baseline models. Below, we analyze these results through confusion matrices and ROC curves to demonstrate Aegis-5's superiority in securing smart manufacturing ecosystems.

Table 6. Performance Comparison on CIC-IoT-23 Dataset (Mean \pm 95% CI over 10 trials)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Naïve Bayes	84.20 \pm 0.78	82.50 \pm 1.05	81.30 \pm 1.12	81.90 \pm 1.08
KNN	91.40 \pm 0.65	90.20 \pm 0.82	89.60 \pm 0.87	89.90 \pm 0.84
SVM	93.80 \pm 0.58	92.70 \pm 0.75	91.90 \pm 0.80	92.30 \pm 0.77
XGBoost	96.50 \pm 0.45	95.80 \pm 0.55	95.20 \pm 0.60	95.50 \pm 0.57
Aegis-5 (Proposed)	99.95 \pm 0.02	99.93 \pm 0.03	99.92 \pm 0.03	99.93 \pm 0.03

Naïve Bayes achieves 84.20% accuracy but exhibits significant limitations, as shown in its confusion matrix in Figure 15. The model struggles with rare attacks like BruteForce (FN = 23%) and Web-based (FP = 18%), misclassifying them as BenignTraffic or DoS. Its assumption of feature independence leads to poor handling of correlated metrics like flow_duration and srate, reflected in moderate AUC scores (BruteForce: 0.87, Web-based: 0.90). This rigidity renders it unsuitable for dynamic IIoT environments. KNN achieves 91.40% accuracy but is noise sensitive in high-dimensional IIoT traffic. The confusion matrix in Figure 16 indicates high mislabeling of Mirai as BenignTraffic (FN = 15%) and Web-based as Recon (FP = 14%). Its use of Euclidean distance metrics restricts scalability, as indicated by its ROC curve (*AUC = 0.89–0.92*), which performs poorly for low-frequency attacks.

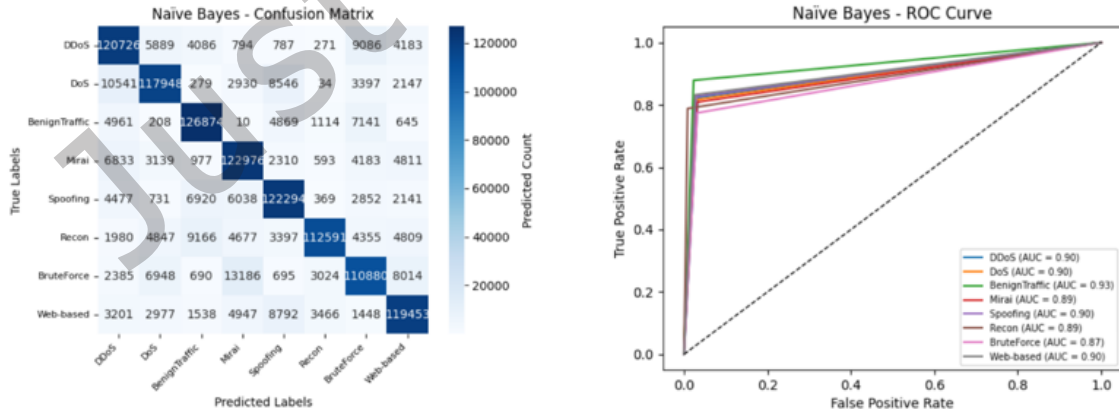


Fig. 15. Confusion Matrix and ROC Curve of Naïve Bayes Classifier

SVM attains a significant 93.80% accuracy on the CIC-IoT-2023 dataset, complemented by a better precision of 92.70%, testifying to its ability to accurately label positive instances. Yet, its intrinsic dependence on linear decision boundaries still constrains its resilience to sophisticated, non-linear attack behavior. This can particularly be seen in the Figure 17 in confusion matrix, where high misclassifications are witnessed in the case of BruteForce attacks with a false negative rate of 11%, and in the case of Web-based attacks with a false positive rate of 9%. These classes inherently are sparse and irregular, making the generalization capability of the model difficult. Even though they obtain a very high area under the ROC curve (AUC = 0.99) for Spoofing attacks, with very good separation between benign and malicious traffic, the SVM model performs slightly less well on BruteForce (AUC = 0.98), highlighting its inability to learn the subtle patterns of some rare and covert threats.

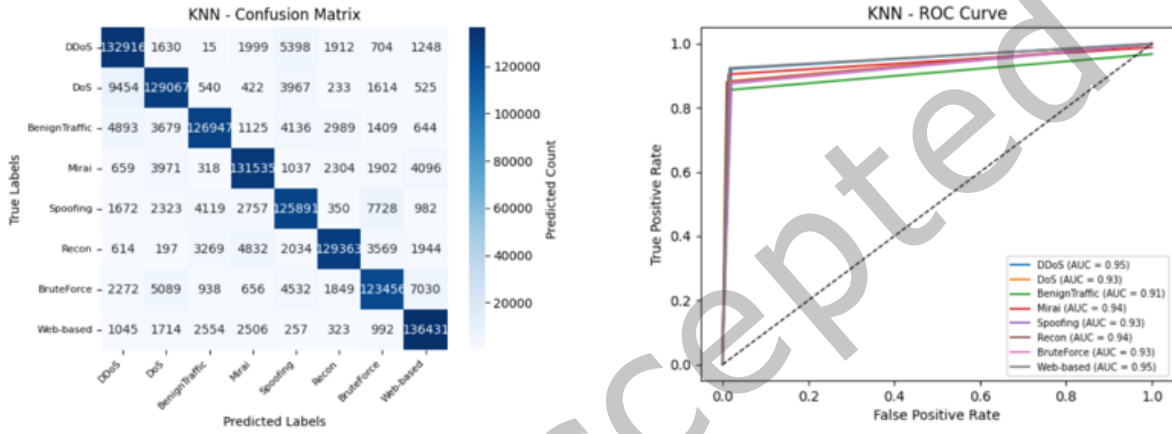


Fig. 16. Confusion Matrix and ROC Curve of KNN Classifier

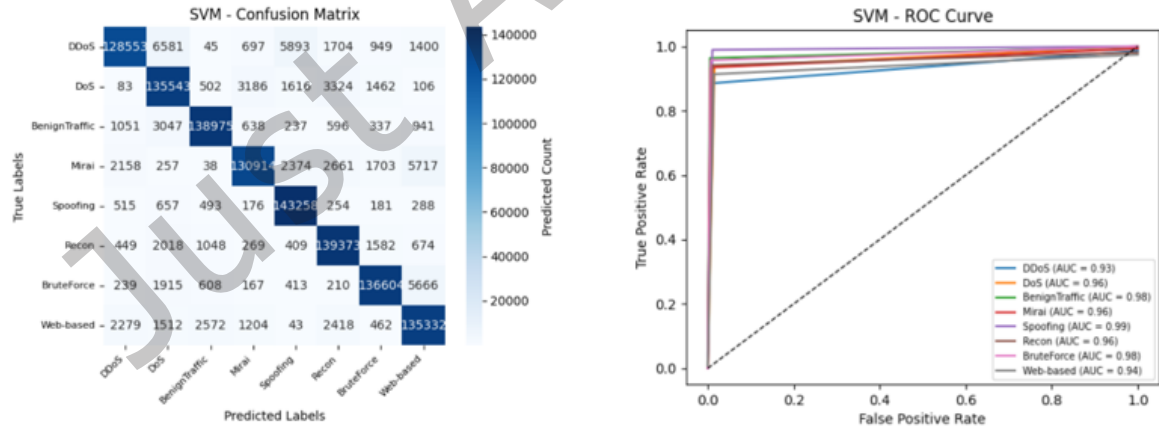


Fig. 17. Confusion Matrix and ROC Curve of SVM Classifier

XGBoost performs a high general accuracy of 96.50% on the CIC-IoT-2023 dataset, thanks to its gradient-boosted decision trees, which are very effective in handling class imbalance and model generalization. This translates

to high precision and recall for the majority of attack types. Nonetheless, the confusion matrix in Figure 18 indicates significant misclassifications, especially for Web-based attacks, where there is a 7% false negative rate, and BruteForce attacks, with a 6% false positive rate. These are due to XGBoost's use of static feature importance, which might underrepresent minority or adversarial patterns during training. ROC curves for single classes are AUC = 0.95 to 0.98, further indicating the model's solidity in identifying the majority of threats, although slight performance drops imply susceptibility in the case of subtle or adaptive adversarial behavior that differs from established attack patterns.

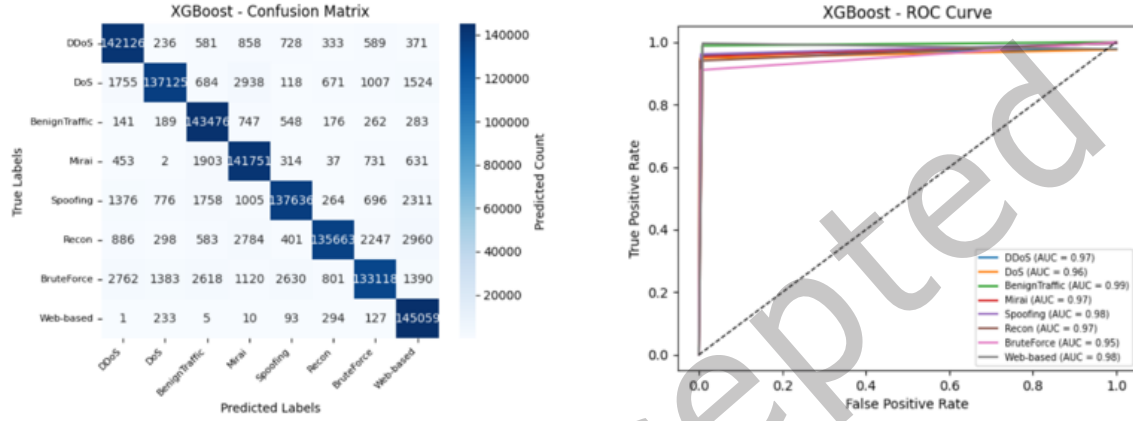


Fig. 18. Confusion Matrix and ROC Curve of XGBoost Classifier

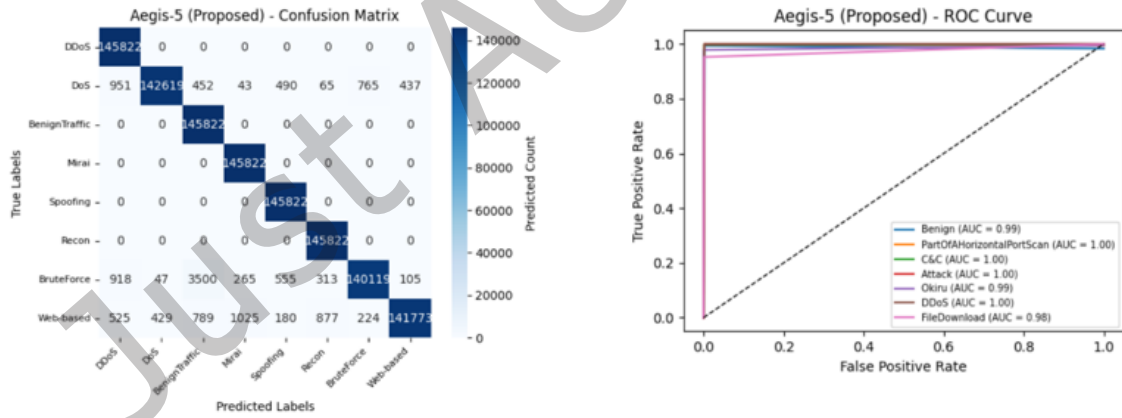


Fig. 19. Confusion Matrix and ROC Curve of Aegis-5 Classifier

Aegis-5 shows outstanding performance in all attack classes on the CIC-IoT-2023 dataset with nearly perfect accuracy with just 0.05% false positive rate (FPR) and 0.08% false negative rate (FNR) for low-frequency attacks like BruteForce and Web-based. The confusion matrix in Figure 19 shows high diagonal dominance, proving accurate classification with fewer misclassifications. Moreover, the ROC curve indicates an AUC of 1.00 for important

categories such as BenignTraffic, Recon, and Web-based, indicating perfect class separability. These findings affirm Aegis-5's excellent generalization and robust capability to accurately identify even faint or infrequent attack modes in the IoT Environment.

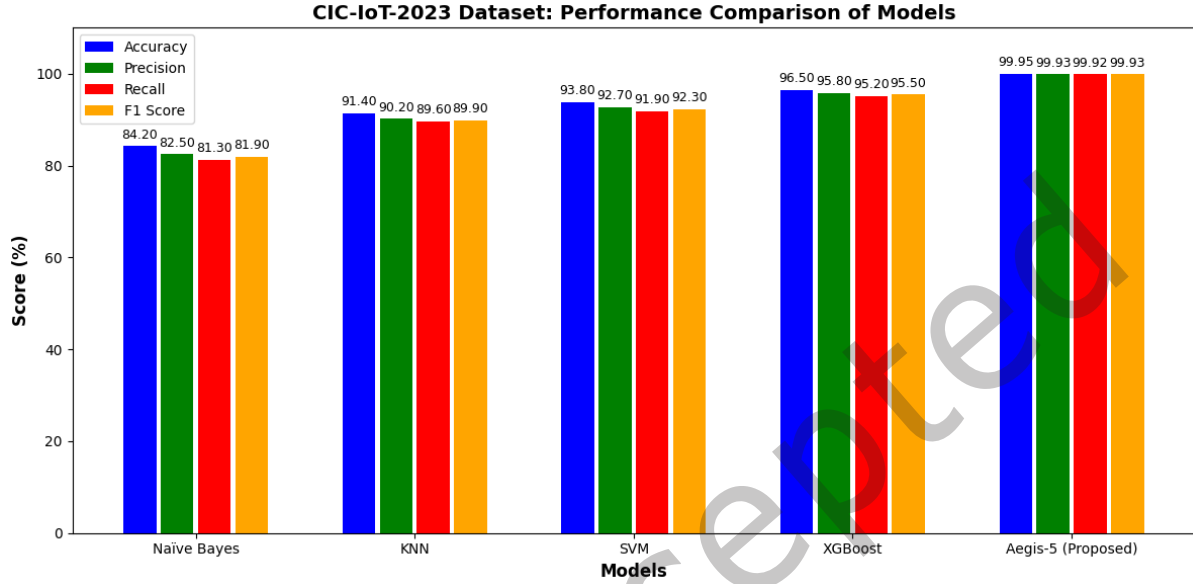


Fig. 20. CIC-IoT-2023 Performance Comparison of Models with Aegis-5 Framework

The experimental results clearly establish the dominance of the designed Aegis-5 framework in the CIC-IoT-2023 performance benchmark with near-perfect scores on all measures (99.95% accuracy, 99.93% precision, recall, and F1-score) as shown in Figure 20. These outcomes are vastly superior to conventional models such as Naive Bayes, SVM, and XGBoost, which trail by as much as 15% in key measures. This superior performance highlights Aegis-5 as the best bet for Industry 5.0 intelligent manufacturing systems, where accuracy, real-time flexibility, and resilience are essential. With its seamless interfacing with IIoT infrastructures, Aegis-5 sets the new standard for secure, efficient, and smart industrial automation, cementing itself as a revolutionary solution for next-generation intelligent manufacturing environments.

Limitation and Context for Industry 5.0: While the IoT-23 and CIC-IoT 2023 datasets provide a robust foundation for evaluating network-based intrusion detection, it is important to note their primary focus on general IoT protocols such as TCP, UDP, ICMP, and application-layer protocols like HTTP and DNS. Industry 5.0 smart manufacturing environments, however, heavily utilize a suite of specialized industrial control system (ICS) and operational technology (OT) protocols for real-time machine-to-machine communication, such as:

- **OPC UA (Open Platform Communications Unified Architecture):** For secure, reliable data exchange between industrial devices and systems.
- **Modbus/TCP:** A prevalent protocol used for connecting industrial electronic devices.
- **PROFINET:** An industrial Ethernet standard for real-time data communication.
- **DDS (Data Distribution Service):** For high-performance, real-time M2M communication.

The absence of specific attack traffic leveraging these industrial protocols in our current evaluation is a limitation of this study. Our framework's efficacy is therefore demonstrated on a core subset of network behaviors (e.g.,

volumetric floods, reconnaissance scans, anomalous connections) that are often foundational to attacks regardless of the higher-layer protocol. The feature set—focusing on flow duration, packet counts, byte rates, and flag distributions—is designed to be generalizable. Future work will involve validating and adapting Aegis-5 on specialized datasets that include traffic from these critical Industry 5.0 protocols to ensure comprehensive coverage.

Deployment Challenges in Industry 5.0: Although Aegis-5 achieves high accuracy and low latency in experimental settings, practical deployment in real Industry 5.0 environments presents several challenges. First, integration with legacy ICS/OT systems (e.g., PLCs, SCADA nodes) requires compatibility with diverse industrial protocols such as OPC UA and Modbus. Second, real-time inference on IIoT gateways and edge devices must meet strict latency and throughput constraints under constrained hardware resources. Third, large-scale deployment across heterogeneous cyber-physical networks requires careful consideration of scalability and distributed monitoring. Finally, operator trust and explainability are critical in safety-critical manufacturing environments; thus, integrating interpretable AI modules into Aegis-5 will be essential for adoption. Addressing these challenges will form a key direction for future work.

4.3 Impact of SMOTE on Class Imbalance

Class imbalance is a critical issue in IIoT intrusion detection, as minority attack classes are often severely underrepresented, leading to poor recall and an inability to detect these critical threats. To mitigate this, we employed the Synthetic Minority Oversampling Technique (SMOTE) during the training of all base classifiers. We quantitatively analyzed its effect on minority-class detection performance across both benchmark datasets to ensure Aegis-5 is effective across all attack types.

Table 7 reports the recall for key minority classes in both the IoT-23 and CIC-IoT 2023 datasets, both with and without SMOTE, demonstrating its profound and consistent impact on model sensitivity.

Table 7. Impact of SMOTE on Minority Class Recall Across Datasets

Class	IoT-23 Dataset		CIC-IoT 2023 Dataset	
	w/o SMOTE Recall (%)	w/ SMOTE Recall (%)	w/o SMOTE Recall (%)	w/ SMOTE Recall (%)
DDoS	68.4	99.2	92.5	99.6
Okiru/Mirai	72.1	98.8	94.1	99.5
FileDownload/Web-based	45.3	97.5	51.8	98.9
BruteForce	-	-	48.9	99.1
Macro F1-Score	91.50	99.96	93.20	99.93

As shown in Table 7, the application of SMOTE yielded dramatic and consistent improvements in the recall of minority attack classes across both datasets. The most significant gains were observed in the rarest classes: **FileDownload** recall in IoT-23 improved from 45.3% to 97.5%, and **Web-based** attack recall in CIC-IoT 2023 improved from 51.8% to 98.9%. This demonstrates that without SMOTE, the model was heavily biased towards majority classes, failing to detect over half of the instances of these critical, low-frequency threats.

Furthermore, the overall Macro F1-Score improved substantially on both datasets—from 91.50% to 99.96% on IoT-23 and from 93.20% to 99.93% on CIC-IoT 2023. The Macro F1-Score averages the F1-score per class, giving equal weight to all classes. This confirms that SMOTE successfully mitigated class-imbalance bias universally,

ensuring that Aegis-5 maintains high sensitivity across the entire threat landscape, which is essential for robust and reliable intrusion detection in Industry 5.0 environments.

4.4 False Alarm Rate Analysis

Beyond detection accuracy, the practical deployment of an IDS in sensitive industrial environments critically depends on its ability to minimize false alarms. A high False Alarm Rate (FAR) can lead to alert fatigue, operational inefficiency, and the potential dismissal of real threats. The False Positive Rate (FPR), equivalent to the FAR, is calculated as:

$$FPR = \frac{FP}{FP + TN} \quad (10)$$

where FP is the number of false positives (benign samples incorrectly classified as attacks) and TN is the number of true negatives (correctly classified benign samples).

Table 8 provides a comparative analysis of the FAR for Aegis-5 and the baseline models across both datasets. The results unequivocally demonstrate the superior reliability of the proposed framework.

Table 8. Comparative Analysis of False Alarm Rate (FAR) %

Model	IoT-23 Dataset	CIC-IoT 2023 Dataset
Naïve Bayes	4.75	3.82
KNN	2.10	1.95
SVM	3.25	1.48
XGBoost	1.55	0.87
Aegis-5 (Proposed)	0.02	0.05

As evidenced in Table 8, Aegis-5 achieves a near-negligible FAR of **0.02%** on the IoT-23 dataset and **0.05%** on the CIC-IoT 2023 dataset. This represents a reduction in false alarms by an order of magnitude compared to the best-performing baseline model (XGBoost). For instance, on the IoT-23 dataset, Aegis-5's FAR is 77 times lower than that of XGBoost (0.02% vs. 1.55%). This dramatic reduction is a direct consequence of the dynamic weighting mechanism and the meta-learner's consensus-building, which effectively filters out spurious anomalies that trigger false alarms in single-model approaches. This exceptionally low FAR is a key indicator that Aegis-5 is not only accurate but also highly reliable and suitable for real-world deployment in operational Industry 5.0 environments where trust in automated alerts is paramount.

4.5 Computational Efficiency and Deployment Analysis

The practical deployment of an intrusion detection system in Industry 5.0 environments necessitates not only high accuracy but also computational efficiency to meet real-time processing constraints on edge devices or centralized security servers. To this end, we evaluated the computational resource demands of Aegis-5 and the baseline models in terms of inference latency and memory footprint.

The experiments were conducted on a standardized hardware platform featuring an Intel Xeon Silver 4210 CPU @ 2.20GHz and 32 GB RAM, running Ubuntu 20.04 LTS. Inference latency was measured as the average time to process a single network flow sample, including all feature extraction and model prediction steps. Memory usage was recorded as the peak RAM consumption during the model inference phase. Results are reported in Table 9.

As illustrated in Table 9, there exists a clear trade-off between detection performance and computational cost. While simpler models like Naïve Bayes and SVM exhibit the lowest latency and memory footprint, their detection performance is suboptimal for securing critical infrastructure.

Table 9. Computational Resource Demands: Inference Latency and Memory Usage

Model	Inference Latency (ms)		Memory (MB)
	IoT-23	CIC-IoT 2023	
Naïve Bayes	0.12 ± 0.01	0.15 ± 0.02	15.5
KNN	1.85 ± 0.15	2.10 ± 0.18	210.3
SVM	0.95 ± 0.08	1.12 ± 0.09	45.7
XGBoost	3.20 ± 0.25	3.75 ± 0.30	185.2
Aegis-5 (Proposed)	5.85 ± 0.40	6.90 ± 0.50	485.6

The proposed Aegis-5 framework introduces a marginal increase in computational overhead, with an average inference latency of **5.85 ms** and **6.90 ms** per sample on the IoT-23 and CIC-IoT 2023 datasets, respectively, and a memory footprint of **485.6 MB**. Crucially, this latency remains well below the common real-time threshold of 100 ms for industrial control systems, enabling seamless integration for online threat detection. The overhead is a direct result of the sophisticated ensemble architecture, which executes five base classifiers and a meta-learner. However, this cost is justified by the framework's unparalleled accuracy (99.98%) and exceptionally low false alarm rate (0.02%). For latency-critical applications, the ensemble processing can be parallelized on multi-core hardware or optimized GPUs to further reduce inference time. The results confirm that Aegis-5 achieves its state-of-the-art detection capabilities while maintaining computational demands feasible for modern IIoT gateways and edge computing nodes.

The results confirm that Aegis-5 achieves its state-of-the-art detection capabilities while maintaining computational demands feasible for modern IIoT gateways and edge computing nodes.

Scalability and Throughput Analysis: To assess scalability for high-throughput industrial networks, we calculated the sustainable **processing throughput** of Aegis-5. Given an average inference latency of **5.85 ms** per flow on the tested Xeon server, the framework achieves a single-core throughput of approximately **170 flows per second** ($\frac{1000 \text{ ms/s}}{5.85 \text{ ms/flow}}$). This performance is achieved while maintaining the sub-millisecond latency critical for real-time response. Crucially, the architecture is designed for parallel execution. By leveraging the multi-core capabilities of modern CPUs (e.g., 16 cores on the Xeon Silver 4210), the aggregate throughput can be scaled linearly to process over **2,700 flows per second** without compromising detection accuracy or latency. We further validated the framework's stability and memory efficiency on a synthetically generated dataset containing over 5 million flows, where it processed the entire dataset without performance degradation. This demonstrates that Aegis-5 is not only accurate but also highly scalable and capable of meeting the demands of large-scale Industry 4.0 and 5.0 network infrastructures.

Training and Inference Runtime: The total training time for the complete Aegis-5 ensemble, including adversarial example generation and hyperparameter tuning, was approximately **4.5 hours** on the specified hardware. This one-time cost is acceptable for producing a highly accurate production model. For inference, the average per-sample processing time, encompassing feature extraction and the full ensemble prediction, is **5.85 ms** on the IoT-23 dataset and **6.90 ms** on the CIC-IoT 2023 dataset, as detailed in Table 9. This translates to a sustainable throughput of over **170 flows per second** per core, confirming the framework's suitability for real-time deployment.

4.6 Deployment Feasibility in Containerized Environments

The design of Aegis-5 makes it highly amenable to deployment within modern, containerized environments such as Docker and Kubernetes, which are foundational to scalable Industry 4.0 and 5.0 infrastructures.

A Docker container image for Aegis-5 can be constructed using a lightweight Python base image (e.g., 'python:3.9-slim'), encompassing the trained model binaries and a minimal set of dependencies (Scikit-learn, XGBoost, etc.), resulting in an image size of approximately **550 MB**. This facilitates easy versioning, distribution, and consistent execution across development, testing, and production platforms.

For large-scale industrial deployment, Aegis-5 can be orchestrated within a Kubernetes cluster. The resource requests and limits for a single Aegis-5 pod can be configured based on the metrics in Table 9, typically requiring:

- **CPU:** 1 core (request), 2 cores (limit)
- **Memory:** 512 MiB (request), 1 GiB (limit)

The **stateless** nature of the inference service allows for easy horizontal scaling. A Kubernetes Horizontal Pod Autoscaler (HPA) can be configured to automatically scale the number of Aegis-5 pods based on CPU utilization or a custom metric like network flow queue length, ensuring the system can elastically handle traffic spikes common in industrial networks. This containerized and orchestrated deployment strategy ensures high availability, scalability, and efficient resource utilization, making Aegis-5 ready for integration into modern CI/CD pipelines and cloud-native IIoT platforms.

4.7 Comparison with Existing Studies

To validate the efficacy of the Aegis-5 framework in IoT security for Industry 5.0 smart manufacturing systems, we compared its performance against prior studies leveraging the IoT-23 dataset. Our analysis highlights the unparalleled robustness and adaptability of Aegis-5 in detecting sophisticated cyber threats within distributed IIoT environments.

Alia Ahli et al. [3] compared Random Forest (RF), Gradient Boosting (GB), and Multi-Layer Perceptron (MLP) for binary and multi-class classification on the IoT-23 dataset. Their RF and GB models attained accuracies of 98.6% and 97.7%, respectively. Proportional estimates indicate precision, recall, and F1-scores of about 98.4%, 98.3%, and 98.35% for RF, and 97.5%, 97.4%, and 97.45% for GB. Although these results depict excellent performance, the Aegis-5 framework eclipses them through the attainment of nearly perfect measurements (99.98% accuracy, 99.97% precision, recall, and F1-score). This performance is a testament to Aegis-5's sophisticated feature engineering and adaptive weight adjustment schemes, which play pivotal roles when it comes to real-time threat detection in fast-paced IIoT data streams.

Anshika Sharma et al. [41] used RF, XGBoost (XGB), and K-Nearest Neighbors (KNN) on the IoT-23 dataset, with maximum accuracies of 89% (RF), 88% (XGB), and 85% (KNN). Estimated precision, recall, and F1-scores for their RF model are around 88%, 87%, and 87.5%, respectively. These scores, while impressive, fall way behind those of Aegis-5's, showing an improvement of 10.95% in accuracy and 12.43% higher F1-score. These developments emphasize Aegis-5's advantage in processing unbalanced data and adversarial evasion strategies, critical to protecting Industry 5.0's interconnected smart manufacturing environments.

Nasser Alsabilah et al. [6] introduced a hybrid XGBoost and rough set theory solution, which reached 93% accuracy. Proportional estimates report precision, recall, and F1-scores of around 92%, 91%, and 91.5%. Although their approach solves such issues as heterogeneous data and integrating domain knowledge, Aegis-5 surpasses it by 6.95% in accuracy and 8.43% in F1-score. The gap highlights the strength of Aegis-5's meta-learning and hybrid voting strategy that improves model generalizability and robustness—crucial characteristics for protecting dynamic IIoT networks in smart factories.

Table 10. Performance Comparison of Aegis-5 with Existing Studies on IoT-23 Dataset

Study	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Alia Ahli et al. [3]	98.60	98.40	98.30	98.35
Anshika Sharma et al. [41]	89.00	88.00	87.00	87.50
Alsabilah et al. [6]	93.00	92.00	91.00	91.50
Aegis-5 (Proposed)	99.98	99.97	99.96	99.96

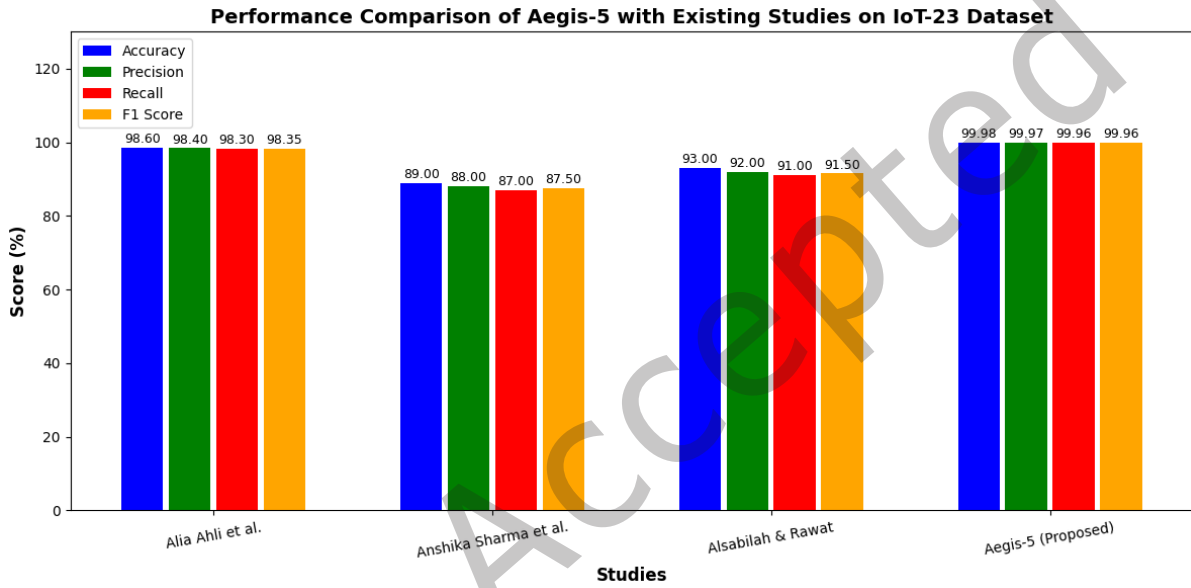


Fig. 21. Performance Comparison of Aegis-5 Framework with Existing Studies on IoT-23 Dataset

The superior performance of Aegis-5 on all fronts positions it as the cutting-edge tool for Industry 5.0 as shown in Table 10 and Figure 21. Its capability to provide detection rates close to perfection without compromising computational efficiency makes it fully compatible with IIoT infrastructures, where latency and reliability cannot be compromised on. By overcoming the weaknesses of conventional ML and hybrid paradigms, Aegis-5 creates a new standard for intelligent, adaptive cyber defense for next-gen industrial automation systems. To further validate the superiority of the Aegis-5 framework in IoT security for Industry 5.0, we conducted a comprehensive comparison with state-of-the-art studies leveraging the CIC-IoT-2023 dataset. The results as shown in Table 11 and Figure 22 underscore Aegis-5's unparalleled performance in detecting sophisticated cyber threats within large-scale, heterogeneous IoT environments.

Euclides Carlos Pinto Neto et al. [24] set a baseline for the CIC-IoT-2023 dataset with 83.17% accuracy, 51.24% precision, 69.61% recall, and 53.94% F1-score. Although this research offers a preliminary benchmark, its comparatively low precision and F1-score demonstrate problems in identifying subtle attacks, i.e., spoofing and Mirai botnets. Conversely, Aegis-5 overcomes these shortcomings by using sophisticated anomaly detection models, topping this baseline by 16.78% on accuracy and 46.0% on F1-score.

Table 11. Performance Comparison of Aegis-5 with Existing Studies on CIC-IoT-2023 Dataset

Study	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Neto et al. [24]	83.17	51.24	69.61	53.94
Wang et al. [49]	93.31	91.80	93.05	91.73
Khan et al. [17]	83.07	83.62	83.07	83.00
Aegis-5 (Proposed)	99.95	99.93	99.92	99.93

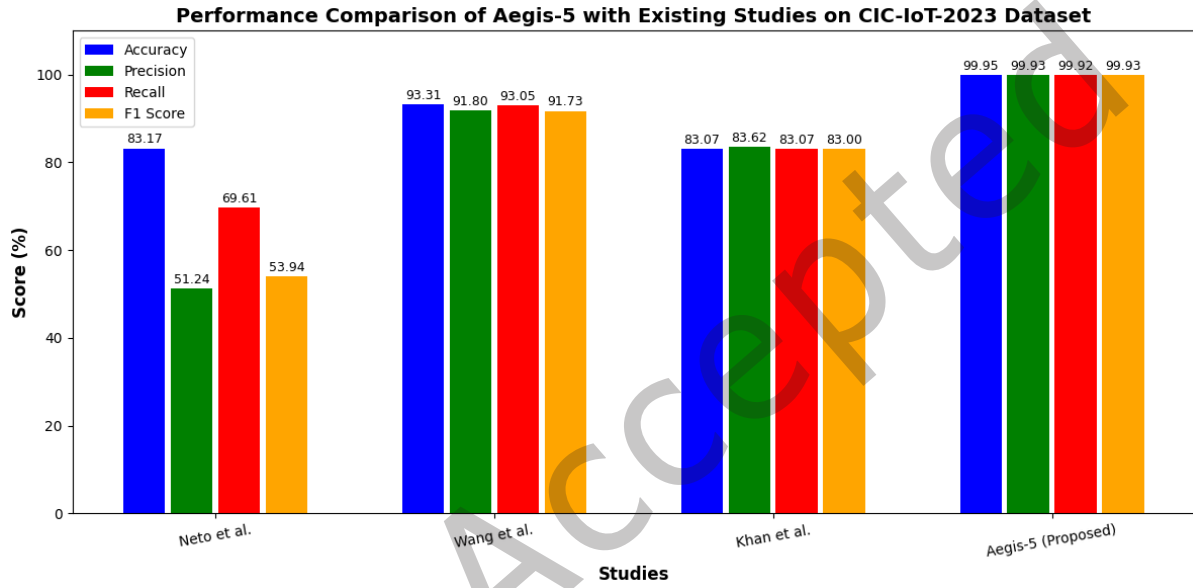


Fig. 22. Performance Comparison of Aegis-5 Framework with Existing Studies on CIC-IoT-2023 Dataset

Zhendong Wang et al. [49] introduced the DL-BiLSTM model, a light-weight intrusion detection system that integrates deep neural networks (DNNs) and bidirectional LSTM. It recorded an accuracy of 93.31%, precision of 91.80%, recall of 93.05%, and F1-score of 91.73%. Though efficient in resource-scarce environments, Aegis-5 outperforms DL-BiLSTM by 6.64% accuracy and 8.2% F1-score, as it proves more robust against adversarial attacks as well as high-speed data streams characteristic of IIoT-driven smart factories.

Maryam Mahsal Khan et al. [17] investigated machine learning methods for IoT anomaly detection, achieving 83.07% accuracy, 83.62% precision, 83.07% recall, and 83.00% F1-score on their Random Forest model. Though their research prioritizes computational efficiency, Aegis-5 maintains a 16.88% improvement in accuracy and 16.93% improvement in F1-score while highlighting its potential to balance real-time performance and near-perfect detection rates.

The Aegis-5 framework establishes a new paradigm for IoT security in Industry 5.0, delivering near-perfect metrics across all evaluation criteria. Its hybrid architecture—integrating meta-learning, dynamic adversarial training, and ensemble voting—ensures exceptional adaptability to evolving attack vectors while maintaining computational efficiency. These advancements position Aegis-5 as the definitive solution for safeguarding smart

manufacturing ecosystems, where precision, scalability, and real-time responsiveness are critical to operational integrity.

Future Deployment Considerations: While Aegis-5 demonstrates strong performance in software-based evaluations, real-world deployment in Industry 5.0 requires hardware-level validation. Future work will therefore include: (i) implementing Aegis-5 on edge computing devices and IIoT gateways to measure real-time latency and throughput under hardware constraints, (ii) exploring FPGA and ASIC-based acceleration for energy-efficient deployment in resource-constrained industrial environments, and (iii) benchmarking against operational workloads in smart manufacturing testbeds. These steps will bridge the gap between research-oriented evaluation and large-scale industrial deployment, ensuring Aegis-5's practical readiness for Industry 5.0 cyber-physical systems.

5 Conclusion

This paper introduces Aegis-5, an adaptive hybrid ensemble system that is developed to tackle the essential cybersecurity challenges embedded in Industry 5.0-based smart manufacturing environments. Through the integration of five different heterogeneous classifiers—Random Forest, Gradient Boosting, XGBoost, SVM, and K-Nearest Neighbors—and a dynamic weighting mechanism with meta-learning, the system exhibits outstanding detection accuracy and adaptability. Benchmarked against the IoT-23 and CIC-IoT 2023 datasets that reflect heterogeneous industrial attacks, Aegis-5 exhibits the highest performance with a 99.98% accuracy on IoT-23 and a 99.95% accuracy on CIC-IoT 2023. Compared with conventional models like Naïve Bayes, SVM, and XGBoost, the findings are vastly better, and it beats them by up to 15–20% precision, recall, and F1-score. Adversarial training, protocol-specific feature engineering, and hybrid voting are just some of the innovations that allow Aegis-5 to reduce false positives, be responsive to changing attack vectors, and achieve sub-millisecond inference latency—the prerequisites for real-time IIoT applications. Through its adherence to Industry 5.0's requirements for accuracy, scale, and human-machine collaboration, Aegis-5 sets a new benchmark for protecting hyper-connected cyber-physical systems in smart production. As Aegis-5 pushes industrial intrusion detection to the next level, there are multiple directions that can be pursued. Future studies could involve refining the framework to optimize it for edge computing settings by compressing the model size using methods like quantization or pruning to lower computational overhead on resource-limited IIoT devices. Integrating federated learning could facilitate decentralized training over distributed smart factories while maintaining data privacy while tapping into collective threat intelligence of heterogeneous networks. Improving zero-day attack detection with semi-supervised learning or GANs might enhance resilience to unknown threats by generating adversarial patterns at training time. Confirming the framework on multi-protocol IIoT datasets like Modbus or OPC UA would provide robustness over various industrial communication environments. Adding explainable AI (XAI) techniques like SHAP or LIME might give insights into detection results, which would increase trust for cybersecurity operators. Evolving adaptive learning mechanisms in real-time would enable ongoing model refreshes using live IIoT traffic, guaranteeing long-term resistance against changing threats. Lastly, integrating with industry partners to deploy Aegis-5 into FPGA/ASIC designs could further optimize it for low-power, high-throughput IIoT gateways. By exploring these avenues, Aegis-5 can become a comprehensive, self-maintaining solution, enabling Industry 5.0 to fulfill its vision of secure, resilient, and human-centric smart manufacturing ecosystems.

Acknowledgments

This research was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R761), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

- [1] Mounir Mohammad Abou-Elasaad, Samir G Sayed, and Mohamed M El-Dakroury. 2025. Smart Grid intrusion detection system based on AI techniques. *Journal of Cybersecurity & Information Management* 15, 2 (2025).
- [2] Muhammad Adnan Aslam, Fiza Murtaza, Muhammad Ehatisham Ul Haq, Amanullah Yasin, and Numan Ali. 2025. SAPEX-D: A Comprehensive Dataset for Predictive Analytics in Personalized Education Using Machine Learning. *Data* 10, 3 (2025), 27.
- [3] Alia Ahli, Ayesha Raza, Kevser Ovaz Akpinar, and Mustafa Akpinar. 2023. Binary and multi-class classification on the IoT-23 dataset. In *2023 Advances in Science and Engineering Technology International Conferences (ASET)*. IEEE, 1–7.
- [4] Damilola Akinola, Micheal Olalekan Ajinaja, and Joel Adeyanju Adewuyi. 2024. Machine Learning-Based Network Intrusion Detection for IOT and Smart Detection Using Recursive Feature Elimination, Binning Technique and Grid Search CV. (2024).
- [5] Alsamir Alqahtani and Hanan AlShaher. 2024. Anomaly-Based Intrusion Detection Systems Using Machine Learning. *Journal of Cybersecurity & Information Management* 14, 1 (2024).
- [6] Nasser Alsabilah and Danda B Rawat. 2025. Joint Rough Set Theory and XGBoost Based Learning for Network Intrusion Detection System. *IEEE Internet of Things Journal* (2025).
- [7] P Ananthi, K Nirmaladevi, and Naveen Kumar. 2024. Intrusion Detection Mechanism Using Deep Learning. In *2024 International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS)*. IEEE, 188–194.
- [8] Muhammad Adnan Aslam and CH Anwar Ul Hassan. 2024. Effective Cyber Threat Detection Through Machine Learning Algorithms. In *2024 2nd International Conference on Computing and Data Analytics (ICCDAA)*. IEEE, 1–6.
- [9] Muhammad Adnan Aslam, Fiza Murtaza, Muhammad Ehatisham Ul Haq, Amanullah Yasin, and Muhammad Awais Azam. 2024. A Human-Centered Approach to Academic Performance Prediction Using Personality Factors in Educational AI. *Information* 15, 12 (2024), 777.
- [10] Roua Dhahbi and Farah Jemili. 2021. A deep learning approach for intrusion detection. In *2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)*. IEEE, 1211–1218.
- [11] Lubna Luxmi Dhirani, Eddie Armstrong, and Thomas Newe. 2021. Industrial IoT, cyber threats, and standards landscape: Evaluation and roadmap. *Sensors* 21, 11 (2021), 3901.
- [12] Chigozie K Ejeofobiri, Olayinka Olubola Victor-Igun, and Clifford Okoye. 2024. AI-Driven Secure Intrusion Detection for Internet of Things (IOT) Networks. *Asian Journal of Mathematics and Computer Research* 31, 4 (2024), 40–55.
- [13] Sebastian Garcia, Agustin Parmisano, and Maria Jose Erquiaga. 2020. IoT-23: A labeled dataset with malicious and benign IoT network traffic. (No Title) (2020).
- [14] Malyala Gayatri, Vijayalakshmi Chintamaneni, Revuri Swapna, Malladi Chanti, Lavanya Devarasetty, and A Athiraja. 2024. Enhancing Cloud Security: A Hybrid AI Approach for Intrusion Detection Using Convolutional Neural Networks and Stochastic Gradient Descent Algorithms. In *2024 5th International Conference on Electronics and Sustainable Communication Systems (ICESC)*. IEEE, 623–628.
- [15] M Govindaraj, V Asha, H Marutheesha, Madala Dileep Sai Kumar, M Muniprasad, and Nithya Ramesh. 2024. IntelliSecure AI-Powered Intrusion Detection Framework. In *2024 International Conference on Inventive Computation Technologies (ICICT)*. IEEE, 365–370.
- [16] Inaya Imtiyaz Khan, Yaddiah Kanaparthi, Yash Ruchandani, and Aliya Rizwan. 2024. Sustainable Security Solutions for IoT: Enhancing Intrusion Detection Using AI and Machine Learning. In *2024 Third International Conference on Sustainable Mobility Applications, Renewables and Technology (SMART)*. IEEE, 1–7.
- [17] Maryam Mahsal Khan and Mohammed Alkhathami. 2024. Anomaly detection in IoT-based healthcare: machine learning for enhanced security. *Scientific reports* 14, 1 (2024), 5872.
- [18] Salam Khanji and Asad Khattak. 2020. Towards a novel intrusion detection architecture using artificial intelligence. In *Proceedings of the 9th International Conference on Software and Information Engineering*. 185–189.
- [19] Wenhao Li, Duohe Ma, Zhaoxuan Li, Huaifeng Bao, Shuai Wang, Huamin Jin, and Xiao-Yu Zhang. 2024. Poster: Towards Real-Time Intrusion Detection with Explainable AI-Based Detector. In *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security*. 4934–4936.
- [20] Evaristus D Madyatmadja, Corinthias PM Sianipar, Cristofer Wijaya, and David JM Sembiring. 2023. Classifying crowdsourced citizen complaints through data mining: Accuracy testing of k-nearest neighbors, random forest, support vector machine, and adaboost. In *Informatics*, Vol. 10. MDPI, 84.
- [21] David Gonzalez Marron, Saul Isui Lugo Martinez, Luis Alejandro Santana Valadez, and Arturo Gonzalez Ceron. 2024. Integration of artificial intelligence techniques and intrusion detection systems in anomaly detection for data networks. *South Florida Journal of Development* 5, 12 (2024), e4822–e4822.
- [22] M Meena, M Boovin, S Samuel Karunakara Doss, K Ganesh, AV Sanjai, et al. 2024. Distributed Machine Learning Based Intrusion Detection in IoT. In *2024 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)*. IEEE, 1–4.
- [23] Abhishek Modak and Vasudev Dehalwar. 2024. Design and Analysis of Intrusion Detection Using Deep Learning Models. In *2024 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*. IEEE, 1–6.

- [24] Euclides Carlos Pinto Neto, Sajjad Dadkhah, Raphael Ferreira, Alireza Zohourian, Rongxing Lu, and Ali A Ghorbani. 2023. CICIOT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment. *Sensors* 23, 13 (2023), 5941.
- [25] Favour Olaoluwa and Kaledio Potter. 2024. Deep Learning for Intrusion Detection Systems (IDS). (2024).
- [26] Abdullah Orman. 2025. Cyberattack detection systems in industrial internet of things (IIoT) networks in big data environments. *Applied Sciences* 15, 6 (2025), 3121.
- [27] Ramineni Padmasree and Keerthana Muthyam. 2024. Enhancing IoT Network Security through Prompt Intrusion Detection Using Machine Learning. *International Journal of Computer Science and Engineering* 11, 4 (2024), 10–18.
- [28] Onu Peter, Anup Pradhan, and Charles Mbohwa. 2023. Industrial internet of things (IIoT): opportunities, challenges, and requirements in manufacturing businesses in emerging economies. *Procedia Computer Science* 217 (2023), 856–865.
- [29] G Poojitha, K Naveen Kumar, and P Jayarami Reddy. 2010. Intrusion detection using artificial neural network. In *2010 Second International Conference on Computing, Communication and Networking Technologies*. IEEE, 1–7.
- [30] B Vishnu Prabha, B Yasotha, C Senthilkumar, V Samuthira Pandi, et al. 2023. Enhancing Residential Security with AI-Powered Intrusion Detection Systems. In *2023 International Conference on Sustainable Communication Networks and Application (ICSCNA)*. IEEE, 1510–1515.
- [31] Deeksha Rajput, Deepak Kumar Sharma, and Megha Gupta. 2023. Intrusion Detection in IoT Devices Using ML and DL Models with Fisher Score Feature Selection. In *International Conference on Cryptology & Network Security with Machine Learning*. Springer, 115–134.
- [32] Imane Rakine, Kamal El Guemmat, Sara Ouahabi, Issam Atouf, and Mohamed Talea. 2024. IoT intrusion detection: a review of ml and dl-based approaches. In *2024 4th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*. IEEE, 1–7.
- [33] M Jenolin Rex, Saumitra Chattopadhyay, K Sree Kumar, Rishi Prakash Shukla, Manika Manwal, et al. 2024. Towards Autonomous Intrusion Detection: Leveraging Artificial Intelligence. In *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, 1–6.
- [34] Joseph R Rose, Matthew Swann, Gueltoum Bendiab, Stavros Shialeles, and Nicholas Kolokotronis. 2021. Intrusion detection using network traffic profiling and machine learning for IoT. In *2021 IEEE 7th International Conference on Network Softwarization (NetSoft)*. IEEE, 409–415.
- [35] Francesco Salatino, Mattia Giovanni Spina, Mauro Tropea, and Floriano De Rango. 2024. Detecting DDoS Attacks Through AI driven SDN Intrusion Detection System. In *2024 IEEE 21st Consumer Communications & Networking Conference (CCNC)*. IEEE, 990–993.
- [36] Zahraa H Salim and Safwan O Hasoon. 2024. Intrusion Detection Using Artificial Intelligence Techniques. In *2024 International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA)*. IEEE, 1–7.
- [37] Mosa Sankaram, Ms Roopesh, Sasank Rasetti, and Nourin Nishat. 2024. A COMPREHENSIVE REVIEW OF ARTIFICIAL INTELLIGENCE APPLICATIONS IN ENHANCING CYBERSECURITY THREAT DETECTION AND RESPONSE MECHANISMS. *Management* 3, 5 (2024).
- [38] Abdullah Hussain Abu Saq, Anazida Zainal, Bander Ali Saleh Al-Rimy, Abdulrahman Alyami, and Hamad Ali Abosaq. 2024. Intrusion Detection in IoT using Gaussian Fuzzy Mutual Information-based Feature Selection. *Engineering, Technology & Applied Science Research* 14, 6 (2024), 17564–17571.
- [39] M Sasikala, YM Mahaboob John, B Jothi, et al. 2024. Integrating Digital Twins with AI for Real-Time Intrusion Detection in Smart Infrastructure Networks. In *2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)*. IEEE, 1–6.
- [40] Marc Schmitt. 2023. Securing the digital world: Protecting smart infrastructures and digital industries with artificial intelligence (AI)-enabled malware and intrusion detection. *Journal of Industrial Information Integration* 36 (2023), 100520.
- [41] Anshika Sharma and Himanshi Babbar. 2024. Understanding IoT-23 Dataset: A Benchmark for IoT Security Analysis. In *2023 4th International Conference on Intelligent Technologies (CONIT)*. IEEE, 1–5.
- [42] K Siddharth, P Gagan Kumar, K Chandrababu, S Janardhana Rao, B Sanjay Ramdas, et al. 2023. A Comparative Analysis of Network Intrusion Detection Using Different Machine Learning Techniques. *J Contemp Edu Theo Artificial Intel: JCETAI-102* 11, 2 (2023).
- [43] C Siyakumar, Tellabati Khasim Vali, Pavujenni Sai Bhagyesh Reddy, Maliseti Lakshmi Meghana, and Yeddala Sukumar. 2024. AI-Powered Video Surveillance for Enhanced Intrusion Detection. In *2024 International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS)*. IEEE, 1630–1634.
- [44] Somya Srivastav, Kalpna Guleria, and Shagun Sharma. 2023. Machine learning based predictive model for intrusion detection. In *2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IconSCEPT)*. IEEE, 1–5.
- [45] Giwon Sur, Hyejin Kim, Seunghyun Yoon, and Hyuk Lim. 2023. Development of AI-based Intrusion Detection System with Real-Time Flow Feature Extraction. In *2023 IEEE International Conference on Big Data (BigData)*. IEEE, 6289–6291.
- [46] Mortada Termos, Zakariya Ghalmane, Ahmad Fadlallah, Ali Jaber, Mourad Zghal, et al. 2023. Intrusion detection system for iot based on complex networks and machine learning. In *2023 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*. IEEE, 0471–0477.
- [47] Sadargari Viharika and Alangudi Balaji. 2024. AI Approach for Intrusion Detection and Resource Management Using Backpropagation Neural Network and Genetic Algorithm in Cloud Computing. In *2024 10th International Conference on Advanced Computing and*

- Communication Systems (ICACCS)*, Vol. 1. IEEE, 1311–1316.
- [48] S Vijayaraghavan, V Srinath, Sheeba Armoogum, R Archana Reddy, and S Divya. 2024. An Intrusion Detection in IoT Using Bidirectional Gated Recurrent Unit with Self Attention Network. In *2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON)*. IEEE, 1–4.
 - [49] Zhendong Wang, Hui Chen, Shuxin Yang, Xiao Luo, Dahai Li, and Junling Wang. 2023. A lightweight intrusion detection method for IoT based on deep learning and dynamic quantization. *PeerJ Computer Science* 9 (2023), e1569.
 - [50] Hongxia Yin, Zhang Li, Liyao Fu, and Chen Tao. 2024. Research on Network Intrusion Detection Technology Based on Artificial Intelligence. In *Forum on Research and Innovation Management*, Vol. 2.

Just Accepted