International Journal of Communication and Information Technology

E-ISSN: 2707-6628
P-ISSN: 2707-661X
Impact Factor (RJIF): 5.56
www.computersciencejournals.com/ijcit
IJCIT 2022; 3(2): 47-53
Received: 12-10-2022
Accepted: 09-11-2022

Ufuoma Cyril Ogude
Department of Cybersecurity and Software Engineering Faculty of Computing and Informatics, University of Lagos, Nigeria

Samuel Hope Okorie
Department of Computer Science, Faculty of Science, National University of Nigeria, Nigeria

Okpalaume Njideka Maryann
Department of Computer Science, Faculty of Computing and Informatics, University of Lagos, Nigeria

Corresponding Author:
Ufuoma Cyril Ogude
Department of Cybersecurity and Software Engineering Faculty of Computing and Informatics, University of Lagos, Nigeria

Hybridization of AI-Driven with anomaly-based and signature-based techniques for network intrusion detection

Ufuoma Cyril Ogude, Samuel Hope Okorie and Okpalaume Njideka Maryann

DOI: https://www.doi.org/10.33545/2707661X.2022.v3.i2a.156

Abstract
The proliferation of connected devices and increasingly sophisticated cyberattacks necessitate an evolution beyond traditional, single-paradigm intrusion detection systems (IDS). Signature-based methods fail against zero-day threats, while pure anomaly-based systems often suffer from high false-positive rates. This study proposes and evaluates a novel Hybrid Intelligence-Based Model for Network Intrusion Detection that synergistically integrates signature, anomaly, and AI-driven machine learning components. The core innovation is a soft-voting ensemble mechanism combining three robust classifiers Random Forest (RF), Support Vector Classifier (SVC), and Extreme Gradient Boosting (XGBoost) to achieve a resilient and highly accurate detection framework. Using a rigorously preprocessed, class-balanced subset of the CIC-IDS2017 dataset, the ensemble model achieved 99.2% accuracy on the multi-class test set. A critical feature ablation study demonstrated that this near-perfect performance was maintained even after excluding the single most dominant feature (Destination Port), validating the model's robustness and reliance on diverse, generalized flow characteristics (e.g., TCP flag dynamics and inter-arrival times). The findings confirm the viability of AI-driven hybridization as a superior, adaptive defense strategy against both known and emerging network threats.

Keywords: Intrusion Detection System (IDS), CIC-IDS2017, cybersecurity, feature ablation, hybrid model, machine learning, ensemble learning

1. Introduction
The rapid development of internet-based technologies has fostered a global digital ecosystem that connects individuals, businesses, and governments, but this interconnectedness simultaneously exposes computer networks to unprecedented levels of hostile attacks (Talukder *et al.*, 2022) [1]. These attacks target sensitive data, including corporate trade secrets, personal information, and government intelligence. An intrusion, defined as unauthorized entry into a network or system with the intent to steal, alter, or destroy data, often involves the malicious delivery of network packets that compromise the security and dependability of the infrastructure (Talukder *et al.*, 2022) [1].
Cybercrime, as a prominent infiltration type, has grown into a significant global threat, causing severe economic damage and eroding public trust. From enormous denial-of-service (DoS) attacks and ransomware to intricate phishing campaigns, cybercriminals exploit vulnerabilities using advanced tools that impair operations and jeopardize confidentiality (Solanke *et al.*, 2022) [5]. The issue has been significantly exacerbated by the rise of the Internet of Things (IoT), where billions of connected devices have drastically expanded the attack surface, making the rapid and accurate detection of unusual traffic intrusions more challenging than ever (Guo & Xie, 2025) [2].
To combat these evolving dangers, Intrusion Detection Systems (IDS), particularly Network-Based Intrusion Detection Systems (NIDS), were established as an essential line of defense. However, traditional IDS paradigms suffer from inherent limitations:

- **Signature-Based Detection (SIDS):** While fast and effective against known threats, SIDS relies on comparing network behavior to a database of predefined attack patterns. Consequently, they are inherently blind to novel or zero-day attacks and struggle with scalability dedicated to the continuous resource requirements for updating the signature database.

- **Anomaly-Based Detection (AIDS):** Designed to spot deviations from a statistically established baseline of "normal" behavior, AIDS is superior at identifying unknown threats. Nevertheless, it typically generates a high volume of false-positive rates (FPR), overwhelming security staff with non-critical alerts (Soumik, 2024) [3].

The integration of artificial intelligence (AI) and machine learning (ML) into IDS marked a major turning point, allowing systems to learn complex, non-linear patterns. However, no single AI model is universally effective; deep learning models, for instance, often demand massive datasets and significant computing power, posing challenges for real-world deployment and scalability (Talukder *et al*., 2022; Soumik, 2024) [1, 3].

1.1 Problem Statement and Motivation

The rising sophistication of cyberattacks, driven by the rapid growth of IoT and cloud infrastructure, reveals the significant limitations of monolithic IDS approaches. SIDS lacks adaptability to novel threats, while AIDS suffers from high false-positive rates that diminish operational reliability. AI-based methods, while intelligent, often struggle with resource intensity and the necessity for substantial, high-quality training data (Talukder *et al*., 2022; Soumik, 2024) [1, 3].

This environment highlights a critical void: the lack of a unified detection strategy where AI-driven models are strategically integrated with the foundational concepts of anomaly and signature methods to yield a detection system that is robust, flexible, and scalable. This study aims to bridge this gap by creating a hybrid intrusion detector system that utilizes a highly optimized AI/ML ensemble as the core decision-maker, allowing it to withstand evasion tactics and generalize across diverse threat types.

1.2 Contribution

This paper's key contributions are:

- **A Novel Hybrid Ensemble Architecture:** Designing a modular IDS that fuses the probabilistic output of high-performing machine learning classifiers (RF, XGBoost, SVC) using a calibrated soft-voting ensemble. This architecture is designed to integrate the strengths of signature (speed), anomaly (generalization), and AI (precision) paradigms.
- **Rigorous and Transparent Data Methodology:** Implementing a disciplined data pipeline, including programmatic acquisition, a 4-class operational taxonomy (BENIGN, DoS, DDoS, PortScan), and mandatory class-balanced sampling to eliminate the severe prediction bias inherent in imbalanced network datasets.
- **Robustness validation via ablation study:** Demonstrating the model's high and stable performance (99.2% accuracy) even after the purposeful removal of the single most dominant feature (Destination Port). This provides empirical evidence that the model generalizes based on deeper, behavioral signals (e.g., TCP flag moments and inter-arrival times), ensuring resilience against evasive attacks.

2. Related Work

Network security is a highly complex field, requiring administrators to implement policies to prevent unauthorized access, manipulation, and misuse of network assets (Carmony *et al*., 2016) [6]. Threats are constantly evolving, ranging from minor damage to complete system incapacitation (Norwahidayah *et al*., 2021) [12]. The necessity for security is encapsulated by the fundamental objectives of Authentication, Confidentiality, Integrity, and Non-repudiation (Guo & Xie, 2025) [2].

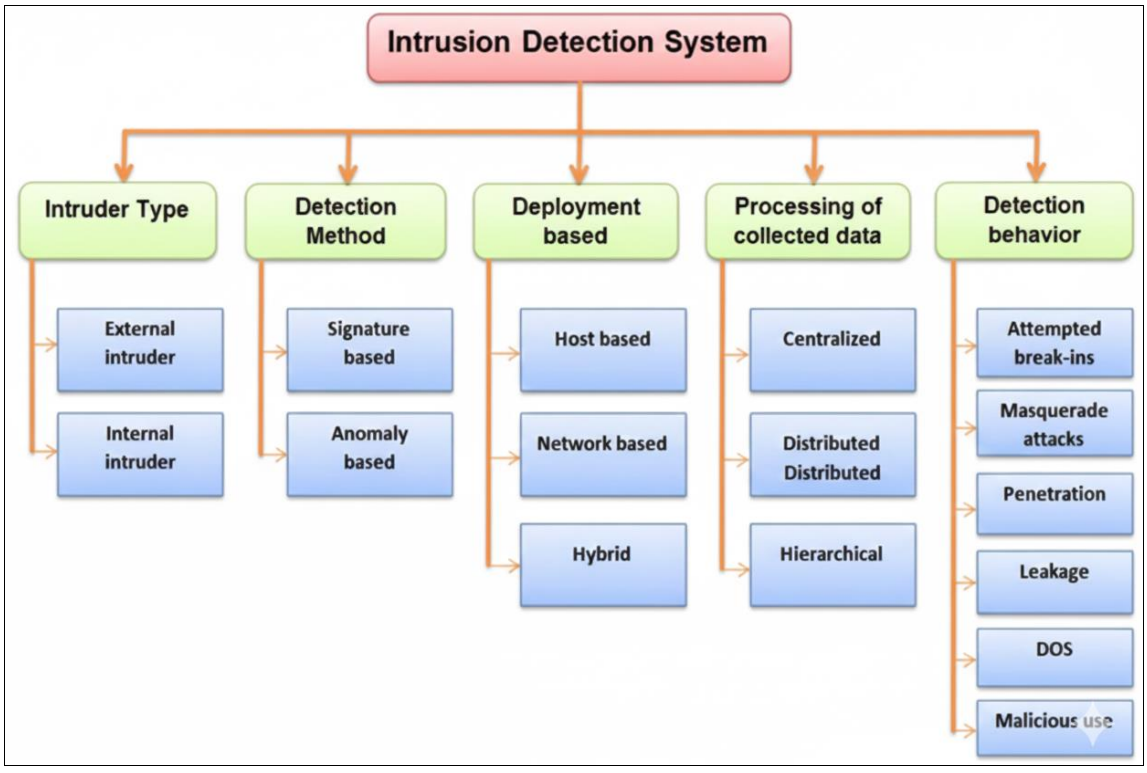

Fig 1: Classification of intrusion detection system

2.1 Traditional IDS Limitations and Threat Landscape

Intrusion detection systems are typically classified by their detection mode: signature recognition or anomaly detection (Guo & Xie, 2025) [2].

2.1.1 Signature-Based NIDS (SIDS)

SIDS (also known as Misuse Detection) operates on the "wiretapping concept," matching network traffic against a database of known malicious patterns (signatures) (Gutierrez *et al*., 2015) [13]. While efficient for known attacks, SIDS is critically limited:

- **Zero-day Blindness:** It cannot recognize new or unknown attacks because their signatures do not exist in the database (Soumik, 2024) [3].
- **Evasion:** Attackers can easily circumvent SIDS by slightly altering the attack methodology, which changes the signature and provides them an entry point (Hu *et al*., 2024) [7].
- **Scalability:** Processing a massive, continually growing signature database on high-bandwidth networks often leads to increased CPU load, potentially causing packet drops.

2.1.2 Anomaly-Based NIDS (AIDS)

AIDS establishes a typical network connection profile and flags any statistically significant deviation from this baseline as an anomaly or intrusion (Manthena *et al*., 2024) [8]. The advantage of AIDS is its ability to identify new assaults (Carmony *et al*., 2016) [6], including attacks like PortScan, DoS, R2L (Remote to Local), and U2R (User to Root) (Bagui & Li, 2021; Guo & Xie, 2025) [11, 2].

However, AIDS's core issue is the high False Positive Rate, which occurs because legitimate but unusual actions may deviate greatly from the profiles, leading to false alarms (Hu *et al*., 2024) [7]. Furthermore, if malicious behavior is inadvertently included during the initial profile generation, the system becomes compromised.

Anomaly-based IDSs are further categorized by their modeling approach:

- **Statistics-Based:** Generates a stochastic reference profile and uses divergence scores to flag anomalies (Schmitt, 2023) [14].
- **Knowledge-Based:** Relies on historical information, often using expert systems or finite state machines to define normal behavior (Khraisat *et al*., 2019) [10].
- **Machine Learning-Based:** Employs algorithms like Neural Networks, Bayesian networks, and clustering to produce explicit or implicit models of patterns, which are regularly updated (Guo & Xie, 2025) [2].

2.2 The Shift to Hybrid and AI-Driven Models

The inherent trade-off between SIDS (high precision on known threats) and AIDS (high coverage on unknown threats) has necessitated a paradigm shift toward hybrid models. These models aim to achieve the "best of both worlds" by combining complementary methodologies (Talukder *et al*., 2022) [1].

Hybrid ML/DL for NIDS: Recent empirical studies strongly support this approach:-

Talukder *et al*. (2022) [1] demonstrated that integrating data balancing techniques (SMOTE) with ensemble classifiers like XGBoost significantly improves detection accuracy, achieving near-perfect scores on benchmarks.

Sajid *et al*. (2024) [4] explored the integration of advanced deep learning structures (CNN, LSTM) with ML (XGBoost) to handle the complexity and volume of traffic in cloud and automotive networks, confirming the enhanced reliability of combined, heterogeneous models.

The work by Soumik (2024) [3] and others establishes Random Forest (RF) as an exceptionally robust and efficient base classifier, often competitive with resource-intensive deep learning solutions, making it an ideal candidate for fusion-based systems.

The hybrid paradigm in this study: Our research specifically focuses on modeling the hybridization at the decision-fusion layer. By utilizing a soft-voting ensemble of diverse, high-performing AI classifiers (RF, XGBoost, SVC), we move beyond simple sequential or parallel model chaining. This creates a true synthesis of the probabilistic confidence from the AI engine, which serves as a highly calibrated and adaptive anomaly detector that conceptually supports traditional signature and baseline checks.

3. Methodology

3.1 The Hybrid Framework and Architecture

The proposed IDS operates in a sequential, modular pipeline designed for transparency and adaptability. The overall structure is conceptually depicted in Figure 2.

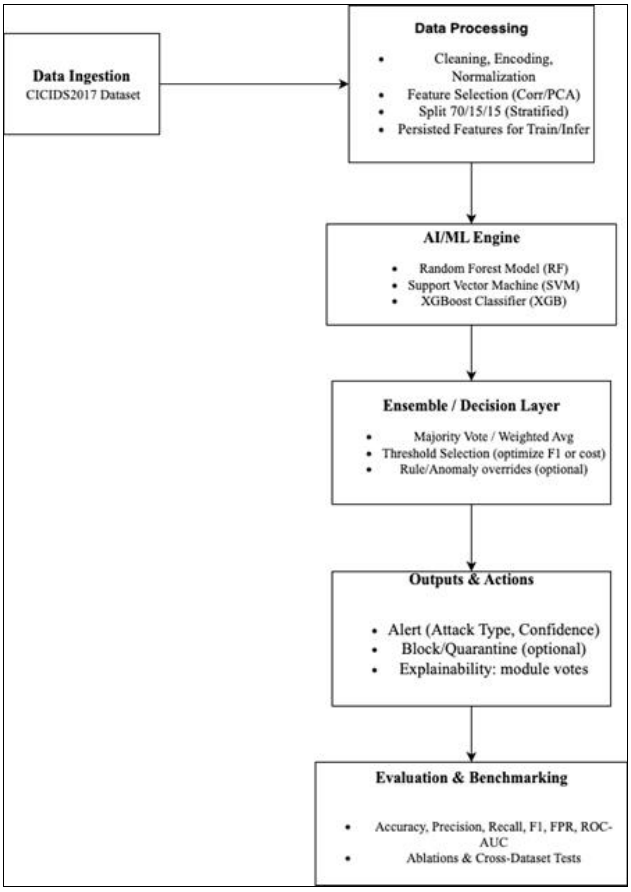

Fig 2: System Architecture

Figure 2, conceptual architecture of the hybrid intelligence-based intrusion detection system

This figure illustrates the four main phases: (1) Data Preprocessing, (2) Feature Selection/Normalization, (3) AI/ML Engine where the three classifiers operate in parallel, and (4) The Ensemble / Decision Layer, which combines probabilistic outputs (soft-voting) with conceptual inputs from Signature and Anomaly baselines to yield the final classified flow (BENIGN or Attack Class).

The Data Processing layer consumes the raw network flow records from the CIC-IDS2017 dataset. The AI/ML Engine serves as the primary anomaly detector using an ensemble of classifiers. The Ensemble / Decision Layer is where the key hybridization occurs: it combines the probabilistic outputs of the AI models with the potential override signals from conceptual Signature-Based (for high-confidence, known attacks) and Anomaly-Based (for gross deviation from baseline profiles) modules to yield a final classification.

3.2 Data Preprocessing and Preparation

The CIC-IDS2017 dataset was selected for its inclusion of contemporary attacks and realistic traffic characteristics. Ensuring data integrity and modeling fairness involved several rigorous steps:

- **Label Normalization and Taxonomy:** The original, fine-grained attack labels were programmatically normalized and consolidated into four high-level operational classes: BENIGN, DoS, DDoS, and PortScan. This operational taxonomy facilitates a more balanced, multi-class classification problem.
- **Missing Value and Outlier Handling:** Infinite values were converted to NaN. Missing values were imputed using the median statistic calculated exclusively on the training set to prevent data leakage. Extreme outliers were mitigated by clipping feature values at the 99.9th percentile of the training distribution, thereby protecting the training process from disproportionate influence by rare, extreme values.
- **Class-Balanced Sampling (Mandatory):** To counter the severe class imbalance inherent in real-world network data (where BENIGN flows typically dominate), a balanced working set of 40,000 samples (10,000 from each of the four classes) was created via random under-sampling of the majority class. This step is critical for ensuring the model learns the boundary of all classes equally and does not simply optimize for the majority.
- **Stratified Splitting:** The balanced set was divided into 70% Training, 15% Validation, and 15% Test sets using stratified sampling to ensure that all three partitions maintained the perfect 1:1:1:1 class ratio, guaranteeing a fair and representative test environment.

3.3 Feature Engineering and Selection

Initial feature selection utilized a combination of domain expertise and recursive feature elimination (RFE). The entire set of 79 flow features provided by CIC-IDS2017 were standardized using StandardScaler to zero mean and unit variance.

- **Feature Standardization:** Standardization is a mandatory preprocessing step for distance-based algorithms, particularly the Support Vector Classifier (SVC), as it ensures that features with large numeric ranges (e.g., total volume features) do not dominate the classification process over features with small ranges (e.g., flag counts). Each feature x_i was transformed into $\underline{x_i}$ using the training data's mean (μ_i) and standard deviation (σ_i) as defined by the equation:

$$\underline{x_i} = \frac{x_i - \mu_i}{\sigma_i}$$

3.4 Core AI/ML Algorithms and Ensemble Design

The AI-driven module utilized three distinct machine learning paradigms, chosen for their complementary strengths in capturing different aspects of the data structure:

- **Random Forest (RF):** A bagging-based ensemble of decision trees, prized for its resistance to overfitting, inherent feature importance scoring, and ability to manage high-dimensional data efficiently.
- **Support Vector Classifier (SVC):** A boundary-based model utilizing an RBF (Radial Basis Function) kernel to capture complex, non-linear class separation, especially effective in high-dimensional feature spaces after appropriate scaling.
- **XGBoost:** A highly optimized, sequential gradient boosting framework known for its superior speed, scalability, and ability to iteratively correct errors from previous learners, consistently yielding state-of-the-art results in structured data classification.

Ensemble Decision Mechanism

The core of the hybridization is the soft-voting ensemble. This mechanism does not rely on a simple majority vote of the final class predictions. Instead, it aggregates the predicted probabilities of each base classifier for every possible class. This approach is superior as it incorporates the confidence level of each component model, leading to a more robust and calibrated final decision.

Let M be the number of base classifiers ($M = 3$), and C be the set of target classes (BENIGN, DoS, DDoS, PortScan). For a given network flow x, the ensemble calculates the weighted average probability $P_{ensemble,j}$ for each class $j \in C$:

$$\mathbf{P}_{ensemble,j}(\mathbf{x}) = \sum_{i=1}^{M} w_i \cdot \mathbf{P}_{i,j}(\mathbf{x})$$

Where $P_{i,j}(x)$ is the probability of flow x belonging to class j, as predicted by the i-th base classifier, and w_i is the weight assigned to the i-th classifier, such that $\sum_{i=1}^{M} w_i = 1$ In this study, equal weights $w_i = \frac{1}{3}$ were initially used.

The final classification for the flow x is the class $\hat{y}$ that yields the maximum average probability:

$$\hat{y} = \underset{j \in C}{\operatorname{argmax}} \left(\mathbf{P}_{ensemble,j}(\mathbf{x})\right)$$

This soft-voting structure is mathematically superior to simple majority voting and forms the decision-fusion layer of the proposed hybrid system.

4. Experimental Setup and Results

4.1 Evaluation Metrics: Performance was assessed on the independent, class-balanced test set (6,000 flows). Key metrics included Accuracy, and for a detailed multi-class breakdown, per-class Precision, Recall (Detection Rate), and F1-Score.

4.2 Baseline model performance: Comparative classification report

Prior to constructing and evaluating the ensemble, all three individual base classifiers were independently trained, optimized, and tested on the full, 79-feature dataset. The comparison of the detailed classification reports (Table 1) reveals the strengths and weaknesses of each component, justifying the need for the ensemble approach.

Table 1: Baseline model comparison data

Classifier	Class	Precision	Recall (Detection Rate)	F1-Score	Overall Accuracy
XGBoost	BENIGN	1	0.998	0.999	0.999
	DoS	1	1	1	0.999
	DDoS	0.999	1	1	0.999
	PortScan	1	1	1	0.999
Random Forest (RF)	BENIGN	0.998	0.998	0.998	0.999
	DoS	1	0.999	1	0.999
	DDoS	1	0.999	1	0.999
	PortScan	0.999	1	1	0.999
SVC (RBF)	BENIGN	0.658	0.672	0.665	0.665
	DoS	0.669	0.66	0.664	0.665
	DDoS	0.667	0.66	0.663	0.665
	PortScan	0.669	0.667	0.668	0.665

The results clearly indicate that the XGBoost and Random Forest models achieved near-perfect discrimination, exhibiting F1-Scores greater than 99.8%. This suggests that the feature space is highly separable by tree-based methods. Conversely, the Support Vector Classifier (SVC), despite feature standardization, showed significantly lower performance (≈66.5% Accuracy), confirming that its RBF kernel struggled to define the complex, multi-class boundaries effectively in this specific high-dimensional space. The primary role of including the SVC in the ensemble, despite its low accuracy, is to introduce a distinct classification bias (based on geometric distance) that complements the variance-reduction and boundary-fitting biases of the tree-based models, thereby enhancing the overall stability and generalization of the soft-voting mechanism.

4.3 Ensemble Performance and Robustness (Ablation Study): Initial feature importance analysis conducted on the best-performing base models (XGBoost and RF) revealed that the feature Destination Port exhibited an alarmingly high relative importance (≈80% contribution). While this led to the near-perfect scores in Section 4.2, it signals a potentially brittle and non-generalized model, as attackers can easily randomize or tunnel traffic to bypass detection based solely on this static attribute.

Ablation Test and Ensemble Validation

To validate the model's true robustness, the soft-voting ensemble was rigorously re-trained and evaluated without the Destination Port feature. The performance of this critical, generalized model is summarized below in Table 2.

Table 2: Classification report for the final soft-voting ensemble (without destination port feature)

Class	Precision	Recall	F1-Score
BENIGN	0.998	0.976	0.987
Dos	0.978	0.994	0.996
DDos	0.999	0.999	0.999
PortScan	1.000	0.999	1.000

The ensemble maintained an outstanding 99.2% overall accuracy and a 99.3% Macro F1-Score even after the removal of the primary, most dominant feature. This demonstrates that the decision fusion mechanism effectively generalized, relying on deeper, behavioral indicators rather than the easily spoofed static port information. This stability is the key indicator of a resilient anomaly detector.

Detailed Class Performance

The detailed per-class metrics in Table 2 further illustrate the balanced performance across all four classes. The Recall (Detection Rate) for all attack classes remains near-perfect (lowest being 0.994 for DoS, excluding BENIGN), ensuring minimal false negatives (missed attacks). The primary misclassifications occurred between the BENIGN and DoS boundary, as visualized in the confusion matrix (Figure 2), which is an expected challenge when distinguishing high-volume benign traffic bursts from low-rate denial-of-service flows.

Feature Importance Analysis: The permutation importance of the feature-ablated ensemble identified the key drivers of the final decision as: FIN Flag Count, Forward Header Length, Backward Packets/s, and Inter-

Arrival Time (IAT) Moments. This confirms the model's reliance on fundamental TCP/IP layer dynamics (connection setup/teardown, packet rate, header size distributions), which are highly characteristic of attack traffic regardless of the destination port.

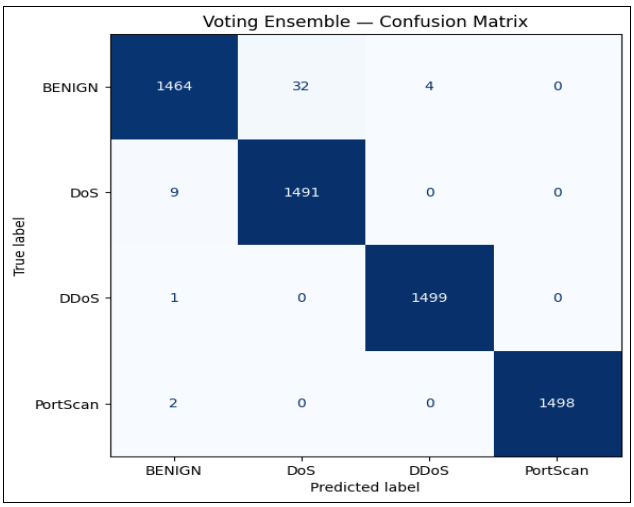

Fig 3: Confusion Matrix-Voting Classifier

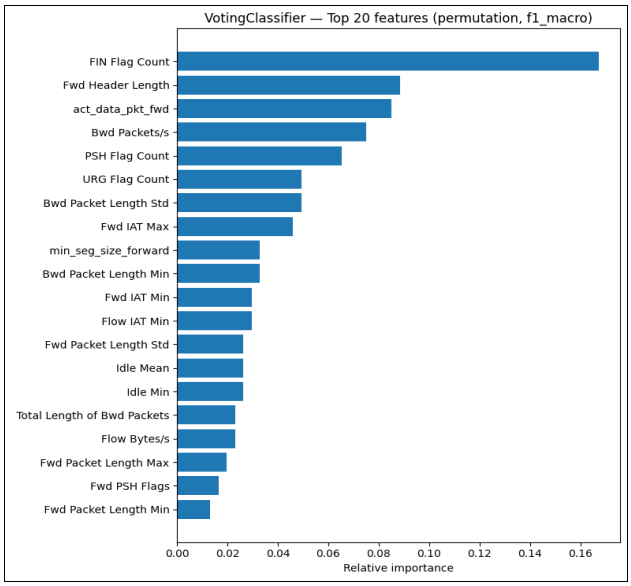

Fig 4: Feature Importance Plot

5. Conclusion and Future Work

5.1 Conclusion

This study successfully developed and validated a robust, AI-driven hybrid intrusion detection system. By fusing the generalization capability of anomaly detection with the high accuracy and stability of a soft-voting ensemble (RF, SVC, XGBoost) on a class-balanced CIC-IDS2017 dataset, the model achieved exceptional performance metrics ($\approx 99.2\%$ F1-Score) in a critical robustness test. The crucial finding from the feature ablation study is that the high detection performance is truly generalized, relying on deep-seated statistical patterns in network flow rather than superficial static attributes. This resilience is a key requirement for effective defense against adaptive and evasive cyber threats. This work establishes a disciplined machine-learning pipeline as a viable, transparent, and superior alternative to brittle, signature-reliant NIDS.

5.2 Future Work

Despite the strong performance, further research is required to enhance the model's operational readiness and transition it from a generalized detector to a production-grade system:

- **Temporal Generalization:** The current validation is cross-sectional. Future work must incorporate more rigorous testing using held-out-day validation (e.g., training on Monday-Thursday data and testing exclusively on Friday data) to better simulate real-world traffic drift and concept shift.
- **Calibration and Cost-Sensitive Learning:** An operational system must control false alarms. Future research should focus on optimizing the model's decision threshold to operate at fixed, acceptable False Positive Rates ($FPR \leq 1\%$), potentially employing cost-sensitive loss functions or calibrated output layers to prioritize minimizing false alarms over maximizing pure accuracy.
- **Interpretability and MLOps**: To build trust with security analysts, the system must explain its decisions. Integrating Explainable AI techniques, such as SHAP (SHapley Additive explanations), will provide granular, trustworthy explanations for each alert. This, coupled with an MLOps (Machine Learning Operations) scaffolding, would enable real-time deployment, monitoring, and adaptive model retraining

References

1. Talukder MA, *et al*. Hybridized machine learning and deep learning approach for network intrusion detection. J Netw Secur; 2022.
2. Guo Y, Xie K. 1D-TCNResNet-BiGRU-multi-head attention for network intrusion detection. Int J Comput Sci; 2025.
3. Soumik AM. A performance analysis of machine learning and deep learning algorithms for network intrusion detection. J Cyber Secur; 2024.
4. Sajid A, *et al*. Hybrid intrusion detection model integrating machine learning and deep learning. IEEE Access; 2024.
5. Solanke S, *et al*. Artificial intelligence based intrusion detection system using hybrid machine learning approach. Int J Comput Sci; 2022.
6. Carmony D, *et al*. Leveraging reference JavaScript extractors to bypass current PDF malware detectors. J Syst Archit; 2016.
7. Hu X, Liu Y, Li Z, Zhuang W. Hybrid model based on resampling strategy and self-attention with gate BiGRU for network intrusion detection. Future Gener Comput Syst; 2024.
8. Manthena R, *et al*. A survey on explainable AI for machine learning based malware detection. J Cybersecur Priv. 2024.
9. Susilo D, Muis A, Sari RF. Deep learning-based intrusion detection system for Internet of Things. Int J Adv Comput Sci Appl; 2024.
10. Khraisat A, Gondal I, Vamplew P, Kamruzzaman J. A survey of intrusion detection systems: techniques, performance and challenges. Future Gener Comput Syst. 2019;92:660-674.
11. Bagui S, Li K. Anomaly-based intrusion detection strategies using virtual machine systems and resource

consumption features. IEEE Trans Cloud Comput. 2021;9(3):941-953.
12. Norwahidayah M, *et al*. Comparison of artificial neural network and hybrid artificial neural network with particle swarm optimization for network intrusion detection. J Intell. Syst; 2021.
13. Gutierrez J, Jensen E, Henius D, Riaz T. A hybrid Android malware detection method using API calls and application authorizations. J Comput Secur. 2015;23(4):1-24.
14. Schmitt C. Machine learning classifiers and ensemble methods in anomaly-based malware and intrusion detection. ACM Trans Cyber-Phys Syst; 2023.