Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control
Journal homepage: http://kinetik.umm.ac.id
ISSN: 2503-2267
Vol. 11, No. 1, February, Pp. 113-122

Weighted ANOVA and mutual information for enhanced intrusion detection system

I Gede Teguh Satya Dharma*[1], I Wayan Rizky Wijaya[1], I Made Agus Oka Gunawan[1], Made Pradnyana Ambara[1]
Politeknik Negeri Bali, Indonesia[1]

Article Info
Keywords:
Intrusion Detection System, Feature Selection, Weighted ANOVA, Mutual Information, Cybersecurity, Machine Learning, Dimensionality Reduction

Article history:
Received: July 28, 2025
Accepted: October 01, 2025
Published: February 01, 2026

Cite:
I Gede Teguh Satya Dharma, I Wayan Rizky Wijaya, I Made Agus Oka Gunawan, and Made Pradnyana Ambara, "Weighted ANOVA and Mutual Information for Enhanced Intrusion Detection System", *KINETIK*, vol. 11, no. 1, Feb. 2026.
https://doi.org/10.22219/kinetik.v11i1.2448

*Corresponding author.
I Gede Teguh Setya Dharma
E-mail address:
teguh@pnb.ac.id

Abstract
The rapid escalation in the sophistication of network attacks has exposed the limitations of traditional Intrusion Detection Systems (IDS). While machine learning has shown great promise in enhancing IDS performance, its success often hinges on the effectiveness of feature selection. Standard feature selection techniques, however, struggle in cybersecurity applications due to the highly imbalanced nature of network traffic datasets. In such settings, minority attack classes—though critical—are often overshadowed by majority classes, leading to reduced detection of rare intrusions. To address this challenge, we propose a hybrid feature selection framework that integrates Analysis of Variance (ANOVA) and Mutual Information (MI) with a novel class-frequency weighting mechanism. This weighting scheme adjusts the relevance score of each feature according to the distribution of classes, ensuring that features associated with rare attacks are more strongly emphasized during the selection process. We evaluate our method on the UNSW-NB15 dataset using a Support Vector Machine classifier. The results show that our approach achieves substantial gains in recall for underrepresented classes while simultaneously reducing feature dimensionality and maintaining efficiency. By improving the visibility of features tied to minority attacks, the proposed framework provides a more balanced and reliable solution for modern IDS. This contribution advances the detection of rare but impactful threats and highlights a scalable pathway for building more resilient cybersecurity defenses.

1. Introduction

The internet has emerged as a fundamental pillar of modern society, facilitating essential infrastructure, enabling global commerce, and supporting uninterrupted communication across geographically dispersed regions [1]. Recent analyses underscore a marked increase in user engagement in the past five years, driven by the proliferation of mobile technologies, scalable cloud platforms, and the widespread availability of high-speed broadband networks [2]. However, this accelerated digital transformation has been accompanied by a corresponding rise of cybersecurity threats. Studies reveal that intrusion attacks constitute a significant proportion of reported cyber-incidents, particularly targeting Internet-of-Things (IoT) systems, critical infrastructure, and cloud-based services [3], [4], [5]. These attacks compromise system availability and integrity, inflict significant economic damage, and undermine public trust.

As adversaries continuously refine their tactics using automated exploitation tools, zero-day vulnerabilities, and advanced social engineering techniques, conventional security mechanisms are becoming increasingly inadequate [6]. This growing complexity highlights the urgent need for adaptive, intelligence-driven security frameworks capable of threat identification and mitigation [7]. Addressing this imperative necessitates a multifaceted approach that incorporates advanced detection algorithms, adaptive access control mechanisms, and robust system architectures tailored for next-generation Internet applications. Recent studies have employed a range of machine learning algorithms to develop intelligent intrusion detection systems, yielding substantial improvements in detection performance over traditional rule-based approaches [8], [9], [10], [11], [12].

Although deep neural networks have demonstrated high accuracy in identifying complex attack patterns, their extensive computational and memory requirements impede real-time operation and scalability in resource-constrained environments [13], [14], [15]. Furthermore, high-dimensional feature spaces introduce significant computational overhead, increase susceptibility to overfitting, and undermine generalization—a challenge commonly referred to as the *curse of dimensionality* [16]. To mitigate these limitations, researchers have investigated dimensionality-reduction and feature-selection techniques such as principal component analysis, heuristic search methods, and evolutionary optimization, effectively pruning redundant or irrelevant features to reduce model complexity while maintaining detection efficacy.

In recent years, feature selection has undergone significant advancements aimed at addressing the challenges posed by high-dimensional data and limited sample sizes. In [17], the authors proposed an efficient network intrusion

detection model for IoT security based on the K-NN classifier and an integrated feature selection framework combining filter, regularization, and embedded paradigms. They employed Principal Component Analysis and SelectKBest to eliminate redundant features via variance-based transformation and univariate statistical ranking. Lasso regression was incorporated to enforce sparsity through L1 regularization, while Random Forest feature importance identifies the most discriminative attributes using impurity reduction metrics. Several other studies have also employed filter-based feature selection methods to enhance the performance of intrusion detection systems.

For instance, researchers in [18], [19], [20] utilized statistical and information-theoretic approaches such as Chi-square and Mutual Information to evaluate the relevance of each feature in relation to the target class. These methods reduce dimensionality while retaining the most informative attributes, ultimately improving classification accuracy and computational efficiency. In addition, tree-based feature selection techniques, which assess feature importance based on decision tree structures, have been applied to identify hierarchical feature relationships and eliminate redundant data. Collectively, these filter-based approaches demonstrate their effectiveness in optimizing feature sets for intrusion detection tasks, particularly when dealing with large and complex network datasets.

In recent years, various hybrid feature selection methods have been proposed for intrusion detection to leverage the complementary strengths of different techniques. A common strategy involves combining filter methods—for example, pairing a statistical metric such as Chi-square with an information-theoretic metric like Information Gain—to capture both linear and non-linear feature relevance [21], [22]. Other approaches adopt a two-stage process, using a computationally efficient filter method for an initial ranking, followed by a wrapper method like Recursive Feature Elimination (RFE) to identify the optimal feature subset from the reduced candidates [23], [24]. Although these hybrid models have advanced the field, their effectiveness remains limited because they typically rely on standard, unweighted metrics that assume equal importance across all data classes.

The reliance on unweighted metrics in existing hybrid methods highlights a critical research gap. In the context of intrusion detection, datasets are notoriously imbalanced, and the assumption of equal class importance can result in models that perform poorly on rare yet critical attack types. Standard implementations of ANOVA and Mutual Information (MI), for example, are susceptible to this issue because they may overlook features that are vital for identifying minority classes. Therefore, the central challenge is not merely to combine linear and non-linear methods but to design a hybrid framework that is inherently sensitive to class imbalance. This study addresses this gap by proposing a novel method that integrates Weighted ANOVA and Weighted Mutual Information. By adjusting feature relevance scores according to class frequencies, the proposed approach constructs a feature set that enhances the detection of underrepresented threats, offering a more robust solution for modern intrusion detection systems.

Weighted ANOVA modifies the computation of the F-statistics by integrating class frequency information, resulting in more reliable identification of discriminative features in skewed datasets [25]. Similarly, Weighted Mutual Information adjusts the influence of each class in entropy-based calculations, allowing greater emphasis on features that are informative for minority classes [26]. Building on these strengths, this study proposes a hybrid filter-based feature selection method that integrates Weighted ANOVA and Weighted Mutual Information into a unified scoring framework. Both metrics are normalized for comparability and then combined through weighting ratios to produce a composite feature ranking. This design directly addresses two common limitations of prior methods: ANOVA alone captures only linear separability, whereas Mutual Information alone is sensitive to sample variance and performs poorly with imbalanced classes. By unifying their complementary strengths and weighting them according to class distributions, the proposed method improves the detection of rare attack categories without sacrificing interpretability or efficiency.

2. Research Methods

This study adopted a four-stage methodological framework encompassing data preprocessing, feature selection, classification, and evaluation. These stages collectively ensured the integrity of the input data and the robustness of the classification outcomes. The UNSW-NB15 dataset [27] was utilized, as it is widely recognized as a modern and comprehensive benchmark for evaluating intrusion detection systems (IDS). Unlike legacy datasets such as KDD99 or NSL-KDD, UNSW-NB15 captured a diverse range of real and synthetic attack types embedded within realistic network traffic flows, making it highly suitable for contemporary cybersecurity research. The CICIDS-2017 dataset was excluded due to several issues related to data mislabeling and artefacts [28].

The data was publicly available and had been pre-divided into training and testing sets without duplication or missing values. Each set contained 36 features, two of which represented the attack category and the binary label indicating normal or malicious traffic. Each feature represents specific aspects of network flow, including basic connection attributes (e.g., protocol type, service), content-based features (e.g., number of failed logins), time-based features (e.g., flow duration), and statistical properties (e.g., packet rate, byte count). These features captured both low-level packet behavior and high-level traffic patterns, making the dataset suitable for evaluating machine-learning-based intrusion detection methods.

The training set has 175,341 rows, while the testing set contains 82,332 rows. An overview of the dataset's features and corresponding descriptions is presented in Table 1. For brevity, only a representative subset of the features is shown, while the full list can be found in the original dataset documentation.

2.1 Data Pre-processing

Features were categorized into six distinct groups based on their semantic properties:
(1) Temporal features, including connection duration (dur), TCP round-trip time components (*tcprtt*, *synack*, *ackdat*), jitter measurements (*sjit*, *djit*), and interpacket timing (*sinpkt*, *dinpkt*);
(2) Packet statistics, encompassing packet counts (*spkts*, *dpkts*) and byte volumes (*sbytes*, *dbytes*);
(3) Protocol information, containing transaction protocol (*proto*), service type (*service*), and connection state (*state*);
(4) TCP-specific metrics, including window advertisements (*swin*, *dwin*) and base sequence numbers (*stcpb*, *dtcpb*);
(5) Statistical measures, covering load rates (*sload*, *dload*), packet loss (*sloss*, *dloss*), and mean packet sizes (*smean*, *dmean*).
(6) Behavioral indicators, including connection counts (*ct_src_dport_ltm*, *ct_dst_sport_ltm*, *ct_flw_http_mthd*, *ct_ftp_cmd*), HTTP transaction depth (*trans_depth*), response body length (*response_body_len*), and binary flags (*is_ftp_login, is_sm_ips_ports*).

Table 1. UNSW-NB15 Dataset Overview (truncated for brevity).

No.	Features	Description
1	*srcip*	Source IP address
2	*sport*	Source port number
3	*dstip*	Destination IP address
4	*proto*	Destination port number
5	*state*	Indicates to the state and its dependent protocol, e.g. *ACC*, *CLO*, *CON*, *ECO*, *ECR*, *FIN*, *INT*, *MAS*, *PAR*, *REQ*, *RST*, *TST*, *TXD*, *URH*, *URN*, and (-) (if not used state)
6	*Dur*	Record total duraction
7	*sbytes*	Source to destination transaction bytes
8	*dybtes*	Destination to source transaction bytes
..	…	…
35	*attack_cat*	The name of each attack category. In this data set, nine categories e.g. Fuzzers, Analysis, Backdoors, DoS Exploits, Generic, Reconnaissance, Shellcode and Worms
36	*label*	0 for normal and 1 for attack records

Temporal features underwent conditional logarithmic transformation based on their distributional properties. Features with absolute skewness exceeding 1.0 were log-transformed after adding a small constant (1e-8) to accomodate zero values, addressing the heavy-tailed distributions commonly found in network timing data. Packet statistics consistently underwent log transformation due to their inherently skewed nature, reflecting the power-law distributions typical of network traffic volumes. Given a feature vector $X = [x_1, x_2, x_3, \dots, x_n]$ the log transformation is formally defined in Equation 1.

$$x_i' = log1p(x_i) = ln(1 + x_i) \quad (1)$$

$$x_i'' = \frac{x_i' - \mu}{\sigma} \quad (2)$$

All numerical features were standardized using zero-mean, unit-variance scaling after their respective transformations, as stated in Equation 2. This normalization ensured equitable contribution across features with disparate scales, ranging from microsecond-level timing measurements to megabyte-scale transfer volumes. The preprocessing pipeline maintained strict temporal separation, with all transformation parameters fitted exclusively on training data and subsequently applied to the validation and test sets to prevent data leakage. Meanwhile, categorical protocol features were transformed using one-hot encoding to convert nominal categories into binary indicator variables. Target labels were preserved without transformation throughout the preprocessing pipeline. The visualization of the data pre-processing pipeline is displayed in Figure 1.

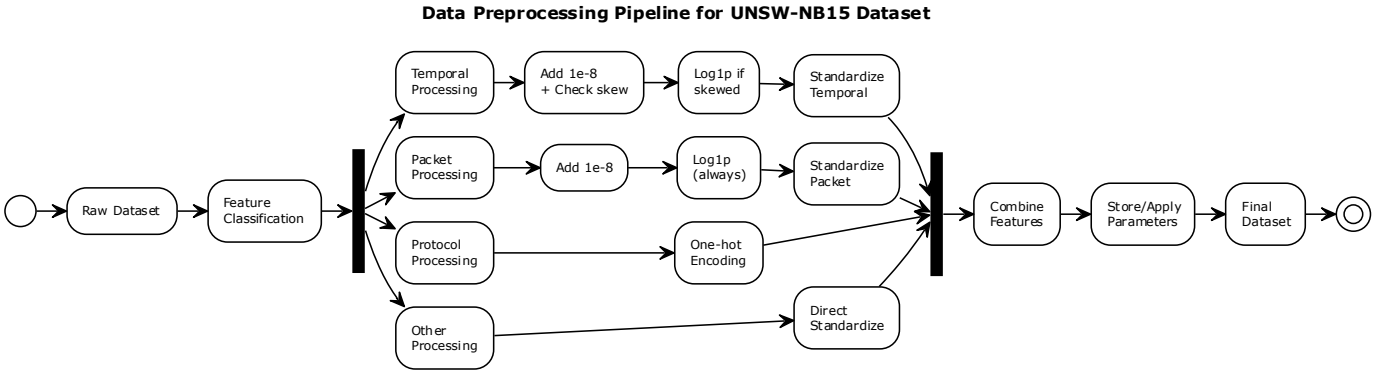

Figure 1. Visualization of the Data Pre-processing Pipeline

2.2 Feature Selection

Conventional feature selection methods often exhibited limitations in simultaneously capturing both linear and non-linear feature-target relationships. Univariate filter approaches (e.g., ANOVA F-test) were effective at identifying linear dependencies but failed to capture complex non-linear associations, whereas (MI was able to capture non-linear patterns but showed higher variance in limited samples. This dichotomy necessitated a unified framework that robustly integrated both modalities while ensuring score comparability across heterogeneous distributions. We introduced a hybrid feature selection that combined the ANOVA F-Test and MI through adaptive score fusion.

In this study, we aimed to identify a parsimonious subset of features from the original space $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ by jointly exploiting class separability and information-theoretic relevance. Let N denote the total number of observations partitioned into C classes of sizes $\{N_1, N_2, \dots, N_C\}$. For each feature x_i, its discriminative power was quantified using the ANOVA F-statistic as defined in Equation 3:

$$S_a(x_i) = \frac{\sum_{c=1}^{C} N_c \left(\overline{x_{i,c}} - \overline{x_i}\right)^2 / (C-1)}{\sum_{c=1}^{C} \sum_{j=1}^{N_c} \left(x_{i,j} - \overline{x_{i,c}}\right)^2 / (N-C)} \tag{3}$$

This statistic measures the extent to which the class-conditional means $\overline{x_{i,c}}$ differ from the global mean $\overline{x_i}$, thereby capturing linear separability across classes. In parallel, nonparametric dependencies were assessed via mutual information, as defined in Equation 4:

$$S_b(x_i) = \sum_{y=1}^{C} \sum_{v} p(x_i = v, y)\, log\, ! \frac{p(x_i=v,y)}{p(x_i=v)\, p(y)} \tag{4}$$

which reflects both linear and nonlinear associations between x_i and the label y. Because $\{S_a(x_i)\}$ and $\{S_b(x_i)\}$ exhibited disparate numerical ranges and distributions, a robust min–max normalization was applied with a zero-range safeguard, as shown in Equation 5:

$$\widehat{S_\alpha}(x_i) = \begin{cases} 0, if\ max_j \widehat{S_\alpha}(x_i) - min_j \widehat{S_\alpha}(x_j) = 0 \\ \frac{\widehat{S_\alpha}(x_i) - min_j \widehat{S_\alpha}(x_j)}{max_j \widehat{S_\alpha}(x_j) - min_j \widehat{S_\alpha}(x_j)}, otherwise \end{cases} \tag{5}$$

This normalization ensures numerical stability even when all raw scores coincide. The normalized measures were fused into a single hybrid relevance score using Equation 6:

$$S_{hyb}(x_i) = w_1 \dot{\widehat{S_\alpha}}(x_i) + w_2 \dot{\widehat{S_\beta}}(x_i), \quad w_1 + w_2 = 1 \tag{6}$$

We began with empirically chosen initial weights w_1 = 0.6 and w_2 = 0.4, which balance sensitivity to between-class variance against information-theoretic content. These weights served as a starting point for further exploration, and the optimal configuration is reported in the *Results and Discussion* section. To derive a definitive feature ordering, features were sorted in descending S_{hyb} values, and each x_i was assigned an integer rank as shown in Equation 7:

$$r(x_i) = 1 + \sum_{j \neq i} 1\left(S_{hyb}(x_j) > S_{hyb}(x_i)\right) \tag{7}$$

such that smaller r-values correspond to higher relevance. The top-k features, with $k = 20$, formed the final selected subset $\{x_{\pi(1)}, \ldots, x_{\pi(k)}\}$, where π orders S_{hyb} in descending magnitude. This hybrid filter-based approach unified linear and nonlinear criteria, enforced scale invariance, and yielded a reproducible, ranked feature set particularly well-suited for high-dimensional intrusion detection tasks.

2.3 Evaluation

The performance of the selected feature set was assessed using four well-established classifiers: Logistic Regression, Decision Tree, Support Vector Machine (SVM), and k-Nearest Neighbor (KNN), in accordance with prior studies [29], [30]. These classifiers were specifically chosen to provide a comprehensive evaluation across a spectrum of machine learning paradigms. Logistic Regression served as a robust linear baseline, valued for its simplicity and interpretability. Support Vector Machine represented high-performance, margin-based model known for its effectiveness in high-dimensional spaces. The Decision Tree offers an interpretable, non-linear approach that models decisions based on explicit rules, while k-Nearest Neighbors provided a non-parametric, instance-based perspective. This diverse selection ensured a thorough assessment of the proposed feature set's performance across different algorithmic assumptions and decision-making processes.

Hyperparameters for each algorithm were optimized via grid search on a stratified validation split, and all experiments were repeated over five random seeds to account for variance in initialization and data sampling. Logistic regression formulated binary classification as a generalized linear model. Given an input feature vector $x \in R^d$ the posterior probability of the positive class was modeled by the logistic function as shown in Equation 8:

$$P(y = 1|x) = \frac{1}{1+exp\left(-(w^T x+b)\right)} \tag{8}$$

where w and b denote the weight vector and bias term, respectively. Parameters were estimated by minimizing the L_2-regularized cross-entropy loss, as shown in Equation 9:

$$L = -\frac{1}{n}\sum_{i=1}^{n}[y_i log\hat{y}_i + (1 - y_i)log(1 - \hat{y}_i)] \; + \; \lambda\|w_2^2\| \tag{9}$$

where $\hat{y}_i = P(y_i|x_i)$, λ is regularization coefficient, and n denotes the number of training samples. The Decision Tree constructed hierarchical partition of the feature space via recursive binary splits. At each node, the feature j and threshold τ were chosen to minimize the weighted Gini impurity, as defined in Equation 10:

$$Gini(D) = 1 - \sum_{k=1}^{K} p_k^2 \tag{10}$$

where p_k is the proportion of samples in node D belonging to class k. Splitting continued until a maximum depth or minimum leaf size was reached, and post-pruning was applied to mitigate overfitting. The last two models, Support Vector Machine (SVM) and k-Nearest Neighbor (KNN), are formally defined in Equations 11 and 12. Support Vector Machine identifies the maximum-margin hyperplane by solving the convex optimization problem, where the slack variables ξ_i allow for misclassification and the regularization parameter C governs the balance between margin width and classification error. Its dual formulation enables kernelization to accommodate non-linear decision boundaries.

$$\min_{w,b,\xi} \frac{1}{2} \; ||w||^2 + C \sum_{i=1}^{n} \xi_i \tag{11}$$

$$d(x, x') = \sqrt{\sum_{j=1}^{d}\left(x_j - x_j'\right)^2} \tag{12}$$

Meanwhile, k-Nearest Neighbor is a nonparametric, instance-based learner that assigns each test sample the majority label among its k-nearest neighbors as measured by the Euclidean distance, eschewing an explicit training phase and extending naturally to multiclass classification. Model performance was evaluated using several metrics: accuracy, precision, recall, F1-score, area under the ROC curve (AUC-ROC), Matthew's correlation coefficient (MCC), false positive rate (FPR), false negative rate (FNR), and specificity. Each metric was computed from the entries of the confusion matrix: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

3. Results and Discussion

This section presents an integrated evaluation of feature selection and classification performance across all examined models. The analysis began by quantifying feature relevance and redundancy within the original dataset to guide subsequent selection procedures. A weighted optimization algorithm was then applied to derive a compact feature

subset that balances interpretability with predictive accuracy. Each model's classification efficacy was assessed using standard metrics.

3.1 Feature Selection Analysis

The comparative evaluation of feature importance scores from weighted ANOVA and Mutual Information (MI) highlights distinct patterns in feature prioritization for intrusion detection. As illustrated in Figure 2, both methods exhibit left-skewed distributions, as features were ranked in descending order of significance. However, ANOVA demonstrated a sharper disparity in scores, with *dbytes* (destination bytes), *state_INT* (connection state), and *dpkts* (destination packets) achieving normalized scores of 1.0, 0.953, and 0.802, respectively.

This pronounced skewness reflected ANOVA's sensitivity to linear variance between attack and normal traffic classes, where features associated with traffic volume and connection state dominated due to their strong discriminative power in linear models [31], [32]. Conversely, MI-based scores, although similarly ranked, displayed a lower standard deviation and a more gradual decline in importance. This behavior indicated MI's ability to capture non-linear dependencies and probabilistic interactions between features. For instance, *sbytes* (source bytes), which was ranked 15th by ANOVA, emerged as the most critical feature in MI analysis. This divergence occurred because MI quantified mutual dependence without assuming linearity, allowing it to identify subtle patterns—such as irregular payload sizes in low-and-slow attacks—that ANOVA may have overlooked due to its reliance on linear separability.

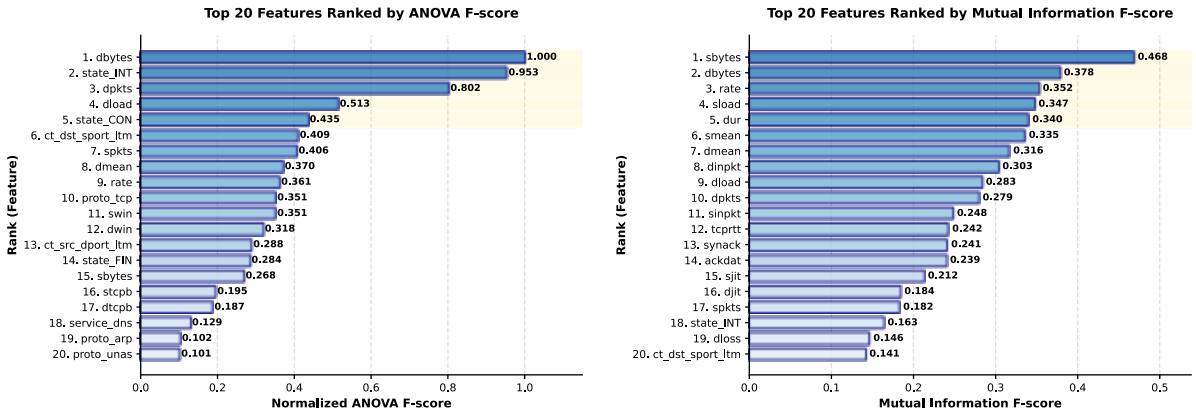

Figure 2. Top 20 Features Rank based on ANOVA and Mutual Information. From left to right: Normalized ANOVA scores and Mutual Information Score

The discrepancy in feature rankings underscored the contextual nature of relevance. ANOVA prioritized features with clear linear discriminability, whereas MI emphasized non-linear associations that were critical for detecting sophisticated attacks. By integrating both approaches, intrusion detection systems achieved a balance between robustness in linearly separable cases with sensitivity to subtle, non-linear anomalies, thereby addressing the evolving landscape of cyber threats.

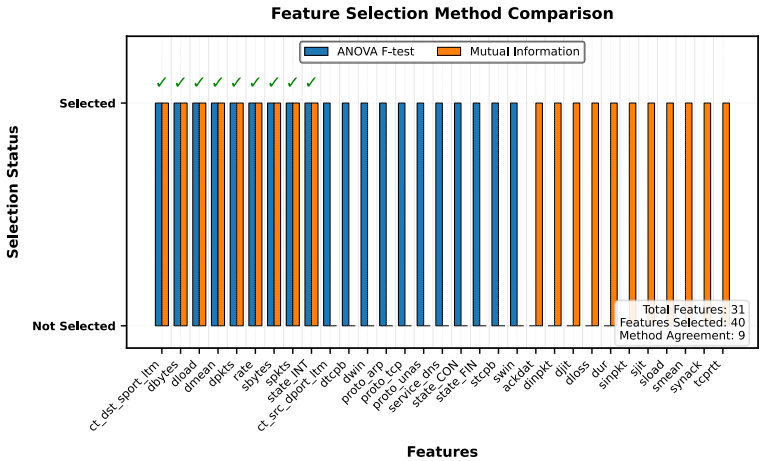

Figure 3. Nine Common Features Selected by ANOVA and Mutual Information

The comparative analysis of the top 20 features selected by weighted ANOVA and Mutual Information (MI), as shown in Figure 3, revealed an intersection of nine common features: *ct_dst_sport_ltm* (count of destination-source port pairs), *dbytes*, *dload*, *dmean*, *dpkts*, *rate*, *sbytes*, *spkts*, and *state_INT*. This overlap highlighted features that consistently exhibited strong discriminative power across both linear (ANOVA) and non-linear (MI) criteria. The shared features primarily represented traffic volume metrics (*dload, dmean, rate*), and connection-state indicators (*state_INT, ct_dst_sport_ltm*). These features were considered critical for intrusion detection due to their universal relevance in characterizing network behavior.

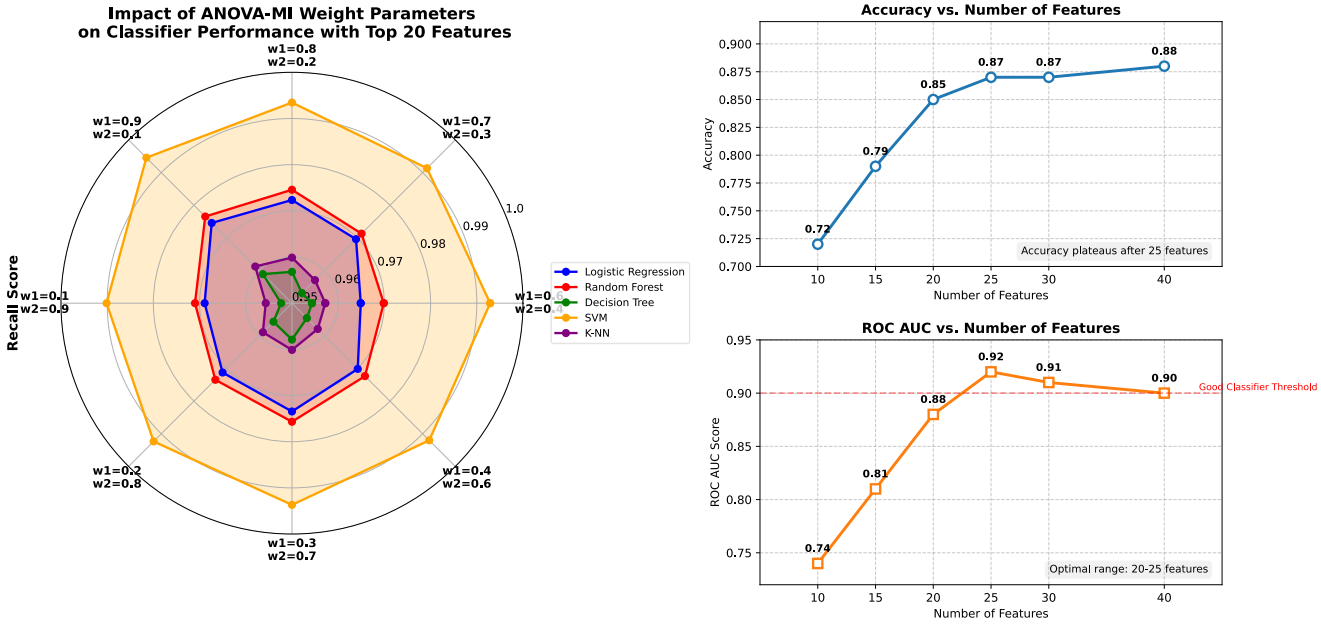

Figure 4. Feature Selection Perfomance Analysis. Left: Classifiers Recall Sensitivity to Weighted ANOVA-MI Parameters using Top 20 Features. Right Column: SVM Classification Performance Metrics versus Feature Count (Top: Accuracy; Bottom: ROC-AUC)

3.2 Classification Performance Evaluation

A sensitivity analysis was conducted to optimize the balance between ANOVA's strength in identifying linear patterns and MI's ability to detect non-linear dependencies. As illustrated in Figure 4 (left panel), the analysis demonstrates that classification performance peaked when the feature set was constructed with a 90% weighting toward ANOVA ($w_1 = 0.9$), enabling the SVM to achieve its highest recall of 99.4%. This pronounced reliance on ANOVA was not a statistical coincidence; rather, it reflected the nature of the cyber threats captured within the UNSW-NB15 dataset. Many of the dataset's attack vectors, including Denial-of-Service (DoS) and Reconnaissance, were fundamentally volumetric. These attacks operated by generating overt statistical anomalies, which manifested as drastic and sustained increases in metrics such as packet counts (*dpkts*), byte volumes (*dbytes*), and connection rates (*rate*). By design, ANOVA was uniquely suited to identify these "loud" and linearly separable patterns, as it precisely measured the statistical variance between the means of normal and attack traffic. While Mutual Information (MI) was theoretically more powerful for detecting stealthy, "low-and-slow" attacks that relied on complex, non-linear feature interactions, its effectiveness was diminished in this context.

The findings indicated that within the mixed-threat environment of UNSW-NB15, the powerful signals generated by volumetric attacks tended to dominate the feature selection process. As a result, features that were highly effective for detecting high-volume attacks overshadowed the subtler indicators required for more sophisticated intrusions. This minimized MI's contribution—not because non-linear patterns did not exist, but because they were rendered less significant for classification when compared to the dataset's dominant, high-variance threats. Therefore, the optimal 0.9:0.1 weighting should be interpreted as a critical diagnostic insight into the dataset itself. It demonstrated that an effective IDS for the UNSW-NB15 environment needed to prioritize the detection of statistically obvious, high-variance threats over more complex, low-amplitude patterns. With this empirically derived weighting, the hybrid method effectively calibrated itself to these specific threat characteristics. The result was a model robustly tuned for volumetric threat detection that also retained a minimal yet strategic capacity for non-linear analysis.

This finding strongly suggested that Mutual Information (MI) contributed minimally to the discriminative power of the feature set, particularly for the minority attack classes. The near-total reliance on the ANOVA component implied a

potential misalignment between the statistical dependencies measured by MI and the complex, high-dimensional nature of network intrusion patterns, where feature relevance was more effectively captured by variance. A comparative evaluation of three machine learning classifiers—Decision Tree, Support Vector Machine, and K-Nearest Neighbor—was performed to scrutinize the feature set's discriminative power. The results highlighted that a surface-level metric such as accuracy, where K-NN appeared superior at 88.3%, was insufficient and masked critical performance dynamics essential for a security context. The superiority of the Support Vector Machine (SVM) became evident when examining the metrics that directly impacted IDS effectiveness. SVM's primary strength lay in its exceptional recall of 99.4%, corresponding to the lowest False Negative Rate (FNR) of 0.007. This result was likely attributable to the SVM's core mechanism—constructing an optimal hyperplane in a high-dimensional space—which made it particularly adept at defining a robust and generalizable decision boundary between legitimate and malicious traffic, even when the patterns were complex and non-linear.

Table 2 Comparative Analysis of Classifier Performance on the UNSW Dataset

Methods	Accuracy (%)	Recall (%)	Precision (%)	F1 (%)	False Positive Rate (FPR)	False Negative Rate (FNR)
Logistic Regression	82.5	93.2	78.5	85.21	0.31	0.068
Decision Tree	86.5	95.5	82.6	88.58	0.25	0.045
Support Vector Machine (SVM	87.3	99.4	82.6	90.22	0.19	0.006
K-Nearest Neighbor (K-NN)	88.3	93.4	85.0	89.00	0.21	0.066

Crucially, this high detection rate did not come at the expense of excessive false alarms. As shown in Table 2, the SVM also achieved the lowest False Positive Rate (FPR) of 0.19, indicating that it was the most effective model at correctly identifying normal traffic without raising unnecessary flags. This dual success in minimizing both FNR and FPR made the SVM stand out. While its precision (82.6%) was slightly lower than K-NN's (85%), the lower FPR suggested that its false positives constituted a smaller fraction of the total normal traffic. In contrast, the K-NN model, while precise, appeared less robust. Its higher FNR (0.066) indicated a potential weakness in identifying more nuanced or novel attacks that did not have close neighbors in the feature space. The Decision Tree, though a competent performer, was likely constrained by its tendency to create axis-parallel splits, which might not have adequately captured the intricate boundaries defining intrusion signatures. Ultimately, the SVM's performance demonstrated a fundamental synergy between its maximum-margin classification approach and the challenge of separating sparse, complex attack vectors from the dense cluster of normal network activity. It navigated the crucial trade-off between security (low FNR) and operational stability (low FPR) more effectively than the alternative models. A comparative analysis was also conducted against several established feature selection methods, with the results presented in Table 3. It was evident that the proposed method yielded a lower overall accuracy (87%) compared to traditional techniques such as Chi Square, which achieved 97.5%.

Table 3 Model Performance Comparison. The best result is annotated with bold fonts

Methods	Accuracy (%)	Recall (%)
Chi Square [33]	**97.5**	92
Info Gain [34]	95.6	90
Wrapper [35]	96.2	91
Weighted MI [36]	92	93
Our work (Weighted ANOVA and MI)	87	**99.4**

However, accuracy was often a deceptive metric in the domain of intrusion detection, where the class distributions were typically imbalanced and the cost of errors was asymmetrical. The paramount objective was to minimize false negatives—a capability directly measured by recall. In this critical aspect, the proposed Weighted ANOVA and MI approach significantly outperformed all other methods, achieving a state-of-the-art recall of 99.4%. This result demonstrated the method's superior ability to construct a feature set highly sensitive to anomalous and malicious patterns. Ultimately, the proposed method represented a strategic trade-off, prioritizing the maximization of threat detection sensitivity over raw predictive accuracy. This choice was deliberate and justified within the security context, where the consequence of failing to detect an attack far outweighed the inconvenience of investigating a false positive. This validated the proposed approach as highly effective for building a robust intrusion detection system. Two primary

limitations were acknowledged. First, the use of static feature snapshots did not capture temporal attack dynamics, suggesting that future work should explore online learning frameworks. Second, although UNSW-NB15 was a comprehensive benchmark, further validation on other datasets was required to confirm the generalizability of the method.

4. Conclusion

This study established that the strategic weighting of ANOVA and mutual information metrics resolved the critical trade-off between linear discriminative capability and non-linear dependency modeling in intrusion detection feature selection. Experimental validation confirmed that a 0.9:0.1 weighting ratio between ANOVA and MI achieved optimal minority-class recall (0.9946 for SVM) by mitigating MI's inherent bias against rare attack vectors while preserving non-linear pattern recognition. The framework maintained an overall accuracy of 87% with only 25 features, demonstrating a 76% reduction dimensionality without performance degradation. However, the pronounced sensitivity to weighting parameters underscored the necessity for domain-informed calibration in class-imbalanced scenarios.

These findings substantiated that optimized weighting was the pivotal factor in enhancing detection robustness. Nevertheless, operational viability required addressing temporal dependency limitations through dynamic feature adaptation to accomodate evolving threat landscapes.

Notation

Fi : Frequency of i-th data.
$S_a(x_i)$: ANOVA features function
$S_b(x_i)$: Mutual Information (MI) features function
$S_{hyb}(x_i)$: Hybrid (Weighted ANOVA and MI) features function
L : Loss function

References

[1] N. Challa, "Unveiling the Shadows: A Comprehensive Exploration of Advanced Persistent Threats (APTs) and Silent Intrusions in Cybersecurity," Journal of Artificial Intelligence & Cloud Computing, pp. 1–5, Dec. 2022. https://doi.org/10.47363/jaicc/2022(1)190

[2] I. A. Kandhro et al., "Detection of Real-Time Malicious Intrusions and Attacks in IoT Empowered Cybersecurity Infrastructures," IEEE Access, vol. 11, pp. 9136–9148, 2023. https://doi.org/10.1109/ACCESS.2023.3238664

[3] B. Alotaibi, "A Survey on Industrial Internet of Things Security: Requirements, Attacks, AI-Based Solutions, and Edge Computing Opportunities," Sensors, vol. 23, no. 17, p. 7470, Aug. 2023. https://doi.org/10.3390/s23177470

[4] N. Singh, R. Buyya, and H. Kim, "Securing Cloud-Based Internet of Things: Challenges and Mitigations," Sensors, vol. 25, no. 1, p. 79, Dec. 2024. https://doi.org/10.3390/s25010079

[5] H. A. Hassan, E. E. Hemdan, W. El-Shafai, M. Shokair, and F. E. A. El-Samie, "Intrusion Detection Systems for the Internet of Thing: A Survey Study," Wirel Pers Commun, vol. 128, no. 4, pp. 2753–2778, Feb. 2023. https://doi.org/10.1007/s11277-022-10069-6

[6] Z. Dai et al., "An intrusion detection model to detect zero-day attacks in unseen data using machine learning," PLoS One, vol. 19, no. 9, Sep. 2024. https://doi.org/10.1371/journal.pone.0308469

[7] K. Bonagiri, P. Krishnamoorthy, V. Keerthiga, D. Kirubakaran, R. David, and B. Nancharaiah, "Cybersecurity With Machine Learning: Implementing AI Algorithms for Intrusion Prevention, Advanced Data Protection, and Real-Time Threat Analysis," in 2025 International Conference on Computational, Communication and Information Technology (ICCCIT), IEEE, Feb. 2025, pp. 292–298. https://doi.org/10.1109/ICCCIT62592.2025.10928115

[8] A. Hussain, A. Khatoon, A. Aslam, and M. A. Khosa, "A Comparative Performance Analysis of Machine Learning Models for Intrusion Detection Classification," Journal of Cyber Security, vol. 6, no. 1, pp. 1–23, 2024. https://doi.org/10.32604/jcs.2023.046915

[9] X. Zhao, "Real-time Application of Intrusion Detection Algorithm Based on Machine Learning in Security System," in 2024 International Conference on Power, Electrical Engineering, Electronics and Control (PEEEC), IEEE, Aug. 2024, pp. 752–757. https://doi.org/10.1109/PEEEC63877.2024.00141

[10] S. Qadir Mohammed and M. A. Hussein, "Performance Analysis of different Machine Learning Models for Intrusion Detection Systems," Journal of Engineering, vol. 28, no. 5, pp. 61–91, May 2022. https://doi.org/10.31026/j.eng.2022.05.05

[11] M. Udurume, V. Shakhov, and I. Koo, "Comparative Evaluation of Network-Based Intrusion Detection: Deep Learning vs Traditional Machine Learning Approach," in 2024 Fifteenth International Conference on Ubiquitous and Future Networks (ICUFN), IEEE, Jul. 2024, pp. 520–525. 10.1109/ICUFN61752.2024.10625037

[12] S. Sreelakshmi, A. A. Babu, C. Lakshmipriya, L. A. Anto Gracious, M. Nalini, and R. Siva Subramanian, "Enhancing Intrusion Detection Systems with Machine Learning," in 2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), IEEE, Oct. 2024, pp. 557–564. https://doi.org/10.1109/ICSSAS64001.2024.10760341

[13] Y. Wang, "Deep Learning-Based Network Intrusion Detection Systems," Applied and Computational Engineering, vol. 109, no. 1, pp. 179–188, Dec. 2024. https://doi.org/10.54254/2755-2721/2024.18104

[14] J. Simioni, E. K. Viegas, A. Santin, and P. Horchulhack, "An Early Exit Deep Neural Network for Fast Inference Intrusion Detection," in Proceedings of the 40th ACM/SIGAPP Symposium on Applied Computing, New York, NY, USA: ACM, Mar. 2025, pp. 730–737. https://doi.org/10.1145/3672608.3707974

[15] R. Picot, F. Gohring de Magalhães, A. Shahnejat Bushehri, M. Ben Atti, G. Nicolescu, and A. Quintero, "Protocol-Agnostic and Packet-Based Intrusion Detection Using a Multi-Layer Deep-Learning Architecture at the Network Edge," IEEE Access, vol. 13, pp. 57867–57877, 2025. https://doi.org/10.1109/ACCESS.2025.3555201

[16] H. Zhang, L. Ge, G. Zhang, J. Fan, D. Li, and C. Xu, "A two-stage intrusion detection method based on light gradient boosting machine and autoencoder," Mathematical Biosciences and Engineering, vol. 20, no. 4, pp. 6966–6992, 2023. https://doi.org/10.3934/mbe.2023301

[17] M. Mohy-eddine, A. Guezzaz, S. Benkirane, and M. Azrour, "An efficient network intrusion detection model for IoT security using K-NN classifier and feature selection," Multimed Tools Appl, vol. 82, no. 15, pp. 23615–23633, Jun. 2023. https://doi.org/10.1007/s11042-023-14795-2

[18] D. Kshirsagar and S. Kumar, “Towards an intrusion detection system for detecting web attacks based on an ensemble of filter feature selection techniques,” Cyber-Physical Systems, vol. 9, no. 3, pp. 244–259, Jul. 2023. https://doi.org/10.1080/23335777.2021.2023651
[19] Q. Liu and Y. Li, “Research on Intrusion Detection Model Based on Filter Feature Selection Algorithm,” in 2024 8th International Conference on Communication and Information Systems (ICCIS), IEEE, Oct. 2024, pp. 120–125. https://doi.org/10.1109/ICCIS63642.2024.10779410
[20] M. A. Umar, Z. Chen, K. Shuaib, and Y. Liu, “Effects of feature selection and normalization on network intrusion detection,” Data Science and Management, vol. 8, no. 1, pp. 23–39, Mar. 2025. https://doi.org/10.1016/j.dsm.2024.08.001
[21] M. . K and N. . S, “An Improved Design for a Cloud Intrusion Detection System Using Hybrid Features Selection Approach with ML Classifier,” International Journal of Innovative Research in Computer and Communication Engineering, vol. 12, no. 10, pp. 11881–11885, Oct. 2024. https://doi.org/10.15680/IJIRCCE.2024.1210063
[22] M. Bakro et al., “An Improved Design for a Cloud Intrusion Detection System Using Hybrid Features Selection Approach With ML Classifier,” IEEE Access, vol. 11, pp. 64228–64247, 2023. https://doi.org/10.1109/ACCESS.2023.3289405
[23] R. Al-Syouf, O. Y. Aljarrah, R. Bani-Hani, and A. Alma’aitah, “Ensemble Machine Learning Models Utilizing a Hybrid Recursive Feature Elimination (RFE) Technique for Detecting GPS Spoofing Attacks Against Unmanned Aerial Vehicles,” Sensors, vol. 25, no. 8, p. 2388, Apr. 2025. https://doi.org/10.3390/s25082388
[24] Y. Yin et al., “IGRF-RFE: a hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset,” J Big Data, vol. 10, no. 1, p. 15, Feb. 2023. https://doi.org/10.1186/s40537-023-00694-8
[25] G. N. N. Barbosa, M. Andreoni, and D. M. F. Mattos, “Optimizing feature selection in intrusion detection systems: Pareto dominance set approaches with mutual information and linear correlation,” Ad Hoc Networks, vol. 159, p. 103485, Jun. 2024. https://doi.org/10.1016/j.adhoc.2024.103485
[26] Y. Zhang, H. Zhang, and B. Zhang, “An Effective Ensemble Automatic Feature Selection Method for Network Intrusion Detection,” Information, vol. 13, no. 7, p. 314, Jun. 2022. https://doi.org/10.3390/info13070314
[27] N. Moustafa and J. Slay, “UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set),” in 2015 Military Communications and Information Systems Conference (MilCIS), IEEE, Nov. 2015, pp. 1–6. https://doi.org/10.1109/MilCIS.2015.7348942
[28] L. Liu, G. Engelen, T. Lynar, D. Essam, and W. Joosen, “Error Prevalence in NIDS datasets: A Case Study on CIC-IDS-2017 and CSE-CIC-IDS-2018,” in 2022 IEEE Conference on Communications and Network Security (CNS), IEEE, Oct. 2022, pp. 254–262. https://doi.org/10.1109/CNS56114.2022.9947235
[29] S. B. Mallampati and H. Seetha, “An Integrated Feature Extraction Based on Principal Components and Deep Auto Encoder with Extra Tree for Intrusion Detection Systems,” Journal of Information & Knowledge Management, vol. 23, no. 01, Feb. 2024. https://doi.org/10.1142/S0219649223500661
[30] Md. B. Pranto, Md. H. A. Ratul, Md. M. Rahman, I. J. Diya, and Z.-B. Zahir, “Performance of Machine Learning Techniques in Anomaly Detection with Basic Feature Selection Strategy - A Network Intrusion Detection System,” Journal of Advances in Information Technology, vol. 13, no. 1, 2022. https://doi.org/10.12720/jait.13.1.36-44
[31] H. A. Al Essa and W. S. Bhaya, “Ensemble learning classifiers hybrid feature selection for enhancing performance of intrusion detection system,” Bulletin of Electrical Engineering and Informatics, vol. 13, no. 1, pp. 665–676, Feb. 2024. https://doi.org/10.11591/eei.v13i1.5844
[32] S. Shakeela, N. S. Shankar, P. M. Reddy, T. K. Tulasi, and M. M. Koneru, “Optimal ensemble learning based on distinctive feature selection by univariate ANOVA-F statistics for IDS,” International Journal of Electronics and Telecommunications, vol. 67, no. 2, pp. 267–275, 2021.https://doi.org/10.24425/ijet.2021.135975
[33] V. S. Bilaskar, S. V. Aradhye, S. S. Shinde, D. D. Kshirsagar, and P. R. Nimbalkar, “An intrusion detection system for industrial IoT using chi-square feature selection,” Journal of Statistics and Management Systems, vol. 27, no. 5, pp. 1021–1031, 2024. https://doi.org/10.47974/JSMS-1303
[34] W. Xu, S. Wang, B. Yan, and Y. He, “Analysis on the Impact of Feature Selection on Cloud Intrusion Detection,” in 2023 4th International Conference on Computer Engineering and Application (ICCEA), IEEE, Apr. 2023, pp. 147–153. https://doi.org/10.1109/ICCEA58433.2023.10135348
[35] Jupriyadi, A. Budiman, E. A. Z. Hamidi, S. Ahdan, and R. M. Negara, “Wrapper-Based Feature Selection to Improve The Accuracy of Intrusion Detection System (IDS),” in 2024 10th International Conference on Wireless and Telematics (ICWT), IEEE, Jul. 2024, pp. 1–5. https://doi.org/10.1109/ICWT62080.2024.10674687
[36] S. Walling and S. Lodh, “Enhancing IoT intrusion detection through machine learning with AN-SFS: a novel approach to high performing adaptive feature selection,” Discover Internet of Things, vol. 4, no. 1, p. 16, Oct. 2024. https://doi.org/10.1007/s43926-024-00074-5