Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15650
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Enhancing Feature-Efficient Network Intrusion Detection Using Gradient Boosting and Chi-Square Selection on NSL-KDD

Gilardinho Javiere Oscoraldo Pedrosa Soares[1)*], Fauzi Adi Rafrastara[2)], Ramadhan Rakhmat Sani[3)]
[1,2)]Informatics Engineering, Faculty of Computer Science, Universitas Dian Nuswantoro, Indonesia
[3)]Information System, Faculty of Computer Science, Universitas Dian Nuswantoro, Indonesia
[1)]111202214732@mhs.dinus.ac.id, [2)]fauziadi@dsn.dinus.ac.id, [3)]ramadhan_rs@dsn.dinus.ac.id

Submitted :Dec 1, 2025 | **Accepted** : Jan 02, 2026 | **Published** : Jan 05, 2026

Abstrak: This study examines the growing complexity of cyber threats that increasingly challenge the effectiveness of traditional Network Intrusion Detection Systems (NIDS). Modern attacks, particularly zero-day intrusions, require detection approaches capable of handling high-dimensional network traffic data. However, existing studies rarely examine the trade-off between feature efficiency and generalization performance in boosting-based NIDS under controlled feature-reduction strategies. Moreover, the role of statistical feature selection in mitigating overfitting in classical boosting models remains underexplored. This study evaluates the performance of NIDS by combining boosting ensemble algorithms, namely AdaBoost, Gradient Boosting, and XGBoost, with filter-based feature selection methods, including Information Gain, Chi-Square, and ReliefF. The NSL-KDD dataset is used as the primary benchmark, with Min–Max normalization applied during preprocessing to ensure numerical feature consistency. Model development is conducted using Orange Data Mining, and performance is assessed through 10-fold cross-validation. Experimental results show that Gradient Boosting achieves the highest baseline accuracy among the evaluated models. Further performance improvements are obtained through feature selection, with the Chi-Square method yielding the best result at 81.2% accuracy using 19 selected features. Information Gain also enhances performance, achieving 80.8% accuracy with 13 features, while ReliefF provides comparatively lower gains. These findings demonstrate that effective feature reduction improves generalization performance, reduces computational complexity, and mitigates overfitting. Overall, the proposed combination of Gradient Boosting and statistical feature selection provides a feature-efficient, generalizable intrusion detection strategy for modern network environments.

Keywords: Chi-Square; Feature Selection; Gradient Boosting; Network Intrusion Detection System; NSL-KDD

INTRODUCTION

Network intrusion refers to unauthorized activities within a computer network that aim to access, alter, steal, or damage data and resources (Mulyanto et al., 2024). To address such threats, Network Intrusion Detection Systems (NIDS) have been developed as hardware- or software-based solutions that monitor network traffic and detect suspicious activities using signature-based and anomaly-based detection strategies (Adzimi et al., 2024; Alsulami et al., 2022). However, the rapid growth and increasing complexity of network traffic have reduced the effectiveness of conventional NIDS in countering advanced cyber threats (Putra & Amarudin, 2025).

Globally, cyberattacks continue to rise in scale and complexity. According to (Microsoft Digital Defense, 2024), around 600 million attacks occur globally each day. Meanwhile (IBM, 2025), the worldwide average cost of a data breach has risen to 4.44 million USD, a 15% increase from the prior year. These statistics reveal the significant economic impact of cyber incidents and highlight the urgent need for intrusion detection systems that are fast, adaptive, and capable of responding to evolving threats.

In Indonesia, similar trends are observed. (BSSN, 2023) identified more than 403.9 million anomalous network traffic events, with Generic Trojan RAT being the most common threat, indicating unauthorized remote

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15650
e-ISSN : 2541-2019
p-ISSN : 2541-044X

access attempts. This increase correlates with the growth of internet users, which reached 221.5 million in 2024 (APJII, 2024). The rapid expansion of digital services, IoT devices, and online platforms has widened the attack surface, making national cybersecurity increasingly vulnerable. As a result, Indonesia requires more intelligent, adaptive intrusion detection solutions.

Although technological advances have been made, many NIDS still rely on signature-based detection, which cannot identify new or zero-day attacks without known patterns (Alsulami et al., 2022; Saputra et al., 2025). Consequently, machine learning approaches are increasingly adopted to detect previously unseen anomalies (Agustina et al., 2024; Natha et al., 2022; Wang et al., 2021). However, high-dimensional datasets such as NSL-KDD introduce challenges due to redundant and irrelevant features that can degrade accuracy, increase computational complexity, and exacerbate overfitting, thereby making effective feature selection essential for improving intrusion detection performance (Bouke et al., 2024; Gupta et al., 2022; Nabi, 2023).

Despite the widespread use of ensemble learning in NIDS, most studies prioritize accuracy improvement without systematically examining the relationship between feature dimensionality and model generalization. This study addresses this gap by proposing a feature-efficient Gradient Boosting approach that leverages Chi-Square–based feature reduction to mitigate overfitting while maintaining generalization performance on NSL-KDD. Recent studies have highlighted ensemble learning and feature selection as effective strategies for improving intrusion detection performance, particularly in reducing overfitting (Jaw & Wang, 2021; Kaushik et al., 2023). However, the effect of the number of selected features on generalization stability in boosting-based models remains underexplored. Accordingly, this work extends existing NIDS literature by evaluating statistical feature selection for boosting-based intrusion detection on high-dimensional datasets.

As a primary point of comparison, (Yuliana et al., 2023) assessed several classification methods, including Decision Tree, KNN, Logistic Regression, and Random Forest, on the NSL-KDD dataset. Their research sought to determine which model most effectively distinguishes between normal and anomalous network traffic using the CRISP-DM methodology, without implementing feature selection. The Decision Tree achieved the highest performance, with 80% on the test set, suggesting a high risk of overfitting due to the performance disparity. A notable limitation of this study is the omission of feature optimization, leaving opportunities for improved generalization unexplored.

A related study by (Chavan & Alone, 2025) employed the CICIDS-2017 dataset, which includes modern and complex attack patterns. Their evaluation of different supervised learning techniques showed that Random Forest exhibited superior performance, achieving 97.20% of accuracy. Although differences in the datasets constrain direct comparisons, the findings demonstrate the significant potential of ensemble methods and support the use of boosting-based approaches in this research.

(Agustina et al., 2024) highlighted the crucial role of feature selection in enhancing intrusion detection performance using the UNSW-NB15 dataset. By reducing 45 features to 4 using Decision Tree and Random Forest before training an ANN model, they achieved 98.3% accuracy. While effective, such drastic reduction risks losing critical information. Therefore, a more gradual, adaptive feature selection strategy as implemented in this research is necessary to maintain both performance and information integrity.

METHOD

All experiments were conducted using Orange Data Mining with default hyperparameters and a fixed random seed of 42 to ensure fair comparison and reproducibility. No hyperparameter tuning was applied to isolate the effect of feature selection on model performance. The research workflow comprised dataset preparation, preprocessing, modeling, feature selection, and performance evaluation, as illustrated in Fig. 1.

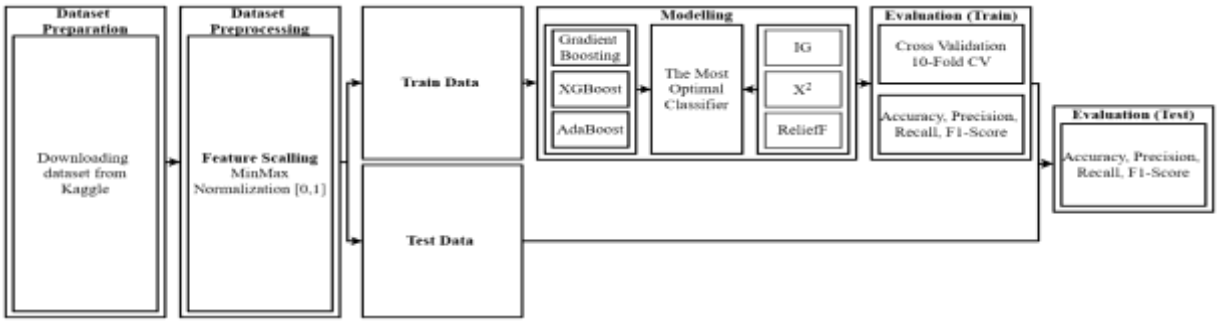

Fig 1. Research Method Flowchart

This study utilized the NSL-KDD dataset obtained from a publicly accessible Kaggle repository, consisting of KDDTrain+ for training and KDDTest+ for independent evaluation. NSL-KDD was selected because it

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15650
e-ISSN : 2541-2019
p-ISSN : 2541-044X

addresses key limitations of the KDD Cup '99 dataset by removing redundant records and reducing severe class imbalance, making it more suitable for machine learning–based intrusion detection (Masoodi et al., 2021; Sahli, 2022). The dataset comprises 125,973 training samples and 22,544 test samples, each with 41 features, including 9 categorical and 32 numerical attributes, as well as a categorical class label. To prevent data leakage, cross-validation was exclusively performed on the KDDTrain+ dataset during the model selection phase. The resulting models were then evaluated on the independent KDDTest+ dataset, ensuring clear separation between the training, validation, and test stages.

Before constructing the model, data preprocessing was carried out to maintain quality and consistency across all input features (Ahmad & Aziz, 2019; Schock et al., 2021). This step included applying Min–Max normalization, which transforms numerical features into a normalized range between 0 and 1 (Ahmed et al., 2022). Normalization is essential to prevent features with larger numerical ranges from having an excessive influence on the learning process, particularly for boosting algorithms that are sensitive to scale variations, while also improving training stability and model convergence (Sujon et al., 2024). Mathematically, Min–Max scaling is expressed in Equation (1), where X' denotes the normalized value, X represents the original feature, and $X_{\min}$ and $X_{\max}$ correspond to the minimum and maximum values of the feature, respectively. Categorical features were encoded using Orange Data Mining defaults, and no missing values or outliers were present in the NSL-KDD dataset.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Table 1 Comparison of Feature Values Before (Left) and After Normalization (Right)

Top	Before Scalling				After Scalling			
	src_bytes	dst_bytes	count	srv_count	src_bytes	dst_bytes	count	srv_count
1	491	0	2	2	3.555806e-07	0	0.00391	0.00391
2	146	0	13	1	1.058e-07	0	0.02544	0.00196
3	0	0	123	6	0	0	0.24070	0.01174
4	232	8153	5	5	1.6812e-07	6.22396e-06	0.00978	0.00978
5	199	420	30	32	1.44207e-07	3.20626e-07	0.05871	0.06262

In this study, three boosting-based ensemble algorithms, Gradient Boosting, XGBoost, and AdaBoost, were employed due to their effectiveness in handling complex datasets and reducing residual errors (Ismanto et al., 2024; Xia et al., 2024). Boosting techniques enhance predictive performance through iterative learning, making them suitable for intrusion detection tasks that require robust pattern recognition. In Gradient Boosting, models are built sequentially, with each iteration correcting the errors of the previous model (Abdullah et al., 2024; Boldini et al., 2023). In this formulation, $F_m(x)$ denotes the model at iteration m, $F_{m-1}(x)$ represents the preceding model, γ_m is the learning rate, and $h_m(x)$ is the weak learner, as expressed in Equation (2).

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad (2)$$

XGBoost, a refined version of Gradient Boosting, incorporates L1/L2 regularization, parallel execution, and sparsity-aware techniques to enhance computational efficiency, scalability, and prevent overfitting (Ardana, 2023; Hakkal & Lahcen, 2024; Shokri et al., 2024). In this formulation, $\mathcal{L}(\phi)$ represents the overall objective function, consisting of two primary elements, the loss function $l(\hat{y}_i, y_i)$, that quantifies the difference between predictions and actual outcomes, and a regularization term $\Omega(f_k)$, that regulates model complexity. Equation (3) presents the XGBoost formulation.

$$\mathcal{L}(\phi) = \sum_{i=1}^{n} l(y^i, yi) + \sum_{k=1}^{K} \Omega(f_k) \quad (3)$$

AdaBoost is a boosting technique that increases the weight of misclassified samples in each iteration, allowing subsequent models to concentrate on harder-to-classify instances and gradually build a more accurate ensemble(Fahrezi et al., 2024; Malashin et al., 2025). In this formulation, $F(x)$ denotes the final model resulting from the combination of all weak learners, $h_m(x)$ represents the weak learner at the m-th iteration, and α_m indicates the learning weight determined based on the accuracy of that weak learner. The AdaBoost formulation is presented in Equation (4).

$$F(x) = \sum_{m=1}^{M} \alpha_h h_m(x) \quad (4)$$

To ensure a more reliable, unbiased, and comprehensive evaluation of the model, this study utilized 10-fold CV (Fonda et al., 2024), a widely used machine learning technique for robust performance assessment. The training dataset is divided into 10 approximately equal parts. The model undergoes 10 iterations, using 9 parts for

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15650
e-ISSN : 2541-2019
p-ISSN : 2541-044X

training and 1 for validation in each iteration, ensuring that every subset function as the validation set a single time. The overall performance is then calculated as the mean across all iterations, providing a thorough measure of model robustness, stability, and reliability, simultaneously mitigating overfitting and promoting better generalization to previously unseen data (Dutschmann et al., 2023; Yuan et al., 2025).

After obtaining the baseline models, feature selection was conducted using three filter-based techniques, Information Gain, Chi-Square (χ^2), and ReliefF is used to select the most significant features and decrease the dimensionality of the data (Hussein et al., 2024; P & S, 2022; Regragui et al., 2024). This step was intended to improve the generalization of the boosting-based models by removing redundant and irrelevant features. Information Gain (IG) quantifies how much a feature reduces entropy or uncertainty about the class label, with features that cause greater reductions being considered more informative (Aqilah Bohani et al., 2024; Azizah et al., 2022; Yang et al., 2024). The method calculates the decrease in class uncertainty produced when a specific feature is incorporated into the model. In this formulation, $IG(Y,X)$ denotes the information gain between feature X and class label Y, $H(Y)$ represents the initial entropy, and $H(Y \mid X)$ indicates the conditional entropy after considering feature X, as shown in Equation (5).

$$IG(Y,X) = H(Y) - H(Y \mid X) \qquad (5)$$

The Chi-Square (χ^2) method evaluates the level of dependence between discretized features and the class label, where a higher χ^2 value indicates significant and relevant relationship for prediction. This technique compares the observed frequency distribution with the expected distribution to assess whether a feature is statistically significantly associated with the target variable (Chairunnisa et al., 2022; Chhaybi & Lazaar, 2025; Ngo et al., 2024). In this formulation, χ^2 represents the Chi-Square statistic indicating the strength of the association between a given feature and its class label, O_i indicates the frequency that was actually observed in the dataset, and E_i indicates the expected frequency assuming independence between the feature and the class label. This formulation is presented in Equation (6).

$$\chi^2 = \sum_{i=1} \frac{(O_i - E_i)^2}{E_i} \qquad (6)$$

ReliefF is a feature selection technique based on instances, which assesses a feature's quality by measuring its capacity to differentiate between neighboring samples of the same class (nearest hit) and those of other classes (nearest miss). The method operates by iteratively adjusting the weight of each feature, increasing the weight of features that effectively separate class boundaries and decreasing the weight of those with low discriminative capability (Freda et al., 2024; Saha & Nandi, 2024; Setiawan, 2023). In this formulation, W[A] represents the weight of feature A, R_i denotes a randomly selected reference sample, H_i is the nearest hit, and M_i is the nearest miss. This formulation is presented in Equation (7). Feature selection was applied once on the training data prior to cross-validation and model evaluation to ensure consistency across experiments and to minimize potential information leakage.

$$W[A] = W[A] - \frac{1}{m}\sum_{i=1}^{m} diff(A, R_i, H_i) + \frac{1}{m}\sum_{i=1}^{m} diff(A, R_i, M_i) \qquad (7)$$

In the final phase of this study, model performance was evaluated using four main metrics such as Accuracy, Precision, Recall, and F1-Score, as these collectively offer a comprehensive assessment of predictive reliability (Rainio et al., 2024; V. Priyalakshmi & Dr. R. Devi, 2022). These metrics were chosen to reflect different aspects of classification performance, particularly for intrusion detection datasets that often have imbalanced class distributions. Accuracy is a performance metric that indicates the proportion of correctly classified test instances, reflecting overall performance across both positive and negative classes. In this context, TP (True Positive) denotes correctly predicted positive instances, TN (True Negative) denotes correctly predicted negative instances, FP (False Positive) denotes incorrectly predicted positive instances, and FN (False Negative) denotes incorrectly predicted negative instances. The formula for calculating accuracy is presented in Equation (8).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (8)$$

Precision is a performance metric that quantifies the proportion of predicted positive instances that are actually positive, making it crucial for reducing false alarms and minimizing misclassification of the positive class. It evaluates the model's capability to generate correct positive predictions while avoiding excessive errors. In this context, TP indicates the number of correctly predicted positive instances, whereas FP (False Positive) represents

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15650
e-ISSN : 2541-2019
p-ISSN : 2541-044X

the number of incorrectly predicted positive instances. The formula for computing precision is provided in Equation (9).

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

Recall is a performance metric that assesses a model's effectiveness in identifying all actual positive instances, which is essential for avoiding missed detections or overlooked positive cases. This measure is significant in intrusion detection systems, where failing to recognize even a single attack can severely compromise network security. In this context, TP (True Positive) denotes the number of correctly predicted positive instances, while FN (False Negative) represents the number of positive cases incorrectly predicted as negative. The formula for recall calculation is provided in Equation (10).

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

F1-Score is a performance metric that calculates the harmonic avergae of precision and recall, a comprehensive and balanced indication of the model's skill in identifying positive samples while minimizing prediction errors. This metric is especially valuable for imbalanced datasets, as it provides a more equitable measure of performance compared to relying solely on accuracy. In this context, Precision denotes the fraction of predicted positive instances that are accurate, and Recall assesses the model's capacity to identify the complete set of actual positive cases. The formula for computing F1-Score is given in Equation (11).

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

RESULT

The initial stage of the experiment was conducted to measure the baseline performance of three boosting-based ensemble algorithms, namely Gradient Boosting, XGBoost, and AdaBoost, using the KDDTest+ test dataset. In this stage, all 41 features were used without any prior feature selection process. The data were first normalized using the Min-Max method to ensure uniform scaling across features. This step aimed to prevent bias toward features with larger value ranges during model training. The preliminary testing results of the three models are presented in Table 2.

Table 2 Model Performance Evaluation Results Without Feature Selection on Test Data

Model	Accuracy	Precision	Recall	F1-Score
Gradient Boosting	79.6%	84.0%	78.3%	78.0%
XGBoost	78.3%	83.6%	79.6%	79.5%
Adaboost	77.7%	81.6%	77.7%	77.7%

As shown in Table 2, the Gradient Boosting model outperforms the other two algorithms, achieving an accuracy of 79.6%. It also achieves the highest precision of 84.0%, demonstrating its strong ability to generate accurate positive predictions. The recall value of 78.3% and F1-Score of 78.0% demonstrate a balanced performance between detection capability and classification accuracy. These results indicate that Gradient Boosting achieves the best overall performance and is therefore selected for the optimization stage.

XGBoost achieves competitive results with an accuracy of 78.3%, which is only slightly lower than Gradient Boosting. The model records the highest recall value of 79.6% and an F1-Score of 79.5%, showing consistent performance in detecting positive classes. Although its precision is marginally lower than Gradient Boosting, the gap in performance between the two models remains relatively small. These results indicate that XGBoost has good generalization capability, although its accuracy remains slightly below that of Gradient Boosting.

AdaBoost shows the lowest performance across all evaluation metrics among the three models. The model records an accuracy of 77.7% and a precision of 81.6%. Its recall and F1-Score values are also both 77.7%. Based on these results, AdaBoost ranks the lowest in the overall model performance comparison. Therefore, this algorithm is less suitable for use with the NSL-KDD dataset.

The three boosting models were additionally evaluated on the training dataset to assess their learning behavior. All models achieved an exceptionally high training accuracy of 99.8%, indicating strong fitting to the training data. However, this performance contrasts sharply with results on the KDDTest+ dataset, revealing a train–test accuracy gap of 20.2%, which indicates severe overfitting in the baseline models. After applying Chi-Square feature selection, this gap was reduced to 18.5%, demonstrating improved generalization capability. These results confirm that feature reduction plays a critical role in mitigating overfitting in boosting-based intrusion detection.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15650
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Once Gradient Boosting was determined to be the top-performing model in the initial evaluation, the next step involved applying three filter-based feature selection techniques, Information Gain (IG), Chi-Square (χ^2), and ReliefF. These techniques were employed to evaluate the impact of individual features with respect to the class label based on their computational characteristics. The feature selection stage aimed to recognize the features with the highest impact on the classification process. Table 3 presents the top five features selected by each method, highlighting the differences in their outputs. This presentation highlights the differences in the outputs of the three techniques and provides insight into which features contribute most to model performance.

Table 3 Comparison of Features and Scores of Feature Selection Methods

Rank	IG		χ^2		ReliefF	
	Feature	Score	Feature	Score	Feature	Score
1	service	0.672	diff_srv_rate	86910.527	service	0.462
2	dst_bytes	0.578	dst_bytes	85616.038	same_srv_rate	0.292
3	src_bytes	0.520	service	82287.635	dst_host_serror_rate	0.267
4	flag	0.519	serror_rate	73702.596	dst_host_count	0.203
5	same_srv_rate	0.464	srv_serror_rate	71790.604	dst_host_same_src_port_rate	0.145

The results reported in Table 3 show that the Information Gain method identifies the feature ‘service’ as the highest-ranked attribute with a score of 0.672, followed by ‘dst_bytes’ and ‘src_bytes’ with scores of 0.578 and 0.520, respectively. The features ‘flag’ and ‘same_srv_rate’ complete the top five with scores of 0.519 and 0.464, indicating the relevance of service type information, data volume, and network interaction patterns in reducing uncertainty toward the target class. These results show that Information Gain tends to prioritize features that contribute substantially to entropy reduction during classification.

For the Chi-Square method, ‘diff_srv_rate’ emerges as the feature with the highest score, reaching 86910.527, followed by ‘dst_bytes’ and ‘service’ with scores of 85616.038 and 82287.635. These features are strongly associated with service patterns and the volume of transferred data, which are key indicators in intrusion-related activities. Furthermore, ‘serror_rate’ and ‘srv_serror_rate’ also appear in significant positions as they reflect standard connection error rates observed in network attacks. These five features are highly relevant for distinguishing normal activities from intrusive behavior.

Meanwhile, the ReliefF method produces a different set of top features, with ‘service’ again ranked first (0.462), followed by ‘same_srv_rate’, ‘dst_host_serror_rate’, ‘dst_host_count’, and ‘dst_host_same_src_port_rate’. These features are closely related to interconnection patterns and host-specific characteristics within network interactions. Although the ranking differs from that of the Chi-Square and Information Gain methods, ReliefF offers an additional perspective by evaluating feature relevance based on proximity to neighboring samples, enabling the detection of indirect relationships among features. However, the resulting scores tend to be smaller and fall within a narrower range than those from the other two methods, indicating lower sensitivity to dominant features.

A comparison across the three methods shows that, despite differences in ranking, there are consistent key features, such as ‘dst_bytes’ and ‘service’, that appear in both the Information Gain and Chi-Square results. This consistency highlights the significant influence of these features in distinguishing attack patterns in the NSL-KDD dataset. In contrast, ReliefF highlights features more closely related to host-level and connection relationships, reflecting a different analytical perspective in assessing feature relevance. Thus, the three feature selection methods complement one another in providing a comprehensive characterization of relevant attributes for intrusion detection.

Subsequent experiments examined the impact of feature selection on the Gradient Boosting model by evaluating different numbers of selected features, ranging from 10 to 20 for each method. The objective was to identify the optimal feature subset that provides the best balance among accuracy, precision, recall, and F1-score. This evaluation also highlights how effectively each feature selection method improves performance compared to using the full feature set. The comparative results of these feature-count variations are summarized in Table 4. In this table, the term “Rank” refers to the number of selected features used in each experiment.

Table 4 Comparison of the Performance of Various Feature Selection Methods

Information Gain				
Rank	Accuracy	Precision	Recall	F1-Score
11	80.0%	84.8%	80.0%	79.8%
12	79.9%	84.7%	79.9%	79.7%
13	80.8%	85.4%	80.8%	80.7%
14	79.6%	84.8%	79.6%	79.5%

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15650
e-ISSN : 2541-2019
p-ISSN : 2541-044X

15	79.5%	84.7%	79.5%	79.3%

Chi Square				
Rank	Accuracy	Precision	Recall	F1-Score
16	79.9%	84.9%	79.9%	79.7%
17	78.9%	84.4%	78.9%	78.8%
18	80.9%	85.5%	80.9%	80.9%
19	81.2%	85.6%	81.2%	81.1%
20	81.0%	85.5%	81.0%	80.9%

ReliefF				
Rank	Accuracy	Precision	Recall	F1-Score
16	77.9%	83.3%	77.9%	77.7%
17	78.6%	83.8%	78.6%	78.5%
18	78.2%	83.9%	78.2%	77.9%
19	79.2%	84.4%	79.2%	79.0%
20	79.4%	84.2%	79.4%	79.2%

Table 4 shows that the Information Gain method produces relatively stable performance on the Gradient Boosting model. The peak accuracy is achieved with 13 features, at 80.8%. In this configuration, precision reaches 85.0%, while recall is 80.8%. The F1-score of 80.7% reflects a harmoniously balanced relationship between detection capability and classification accuracy. These results show that Information Gain is effective in selecting features that meaningfully enhance overall classification performance.

The variation in the number of features also indicates that adding more features does not always lead to better performance. Using 11 to 12 features results in slightly lower accuracy compared to the 13-feature configuration. This suggests an optimal number of features for the Gradient Boosting model. Too few features may cause important information to be lost, whereas too many features may degrade learning quality. Therefore, determining the appropriate number of features is important for maintaining model effectiveness.

For the Chi-Square method, Table 4 shows that it yields the highest performance among Information Gain and ReliefF. The best accuracy is achieved with 19 features, reaching 81.2%. The precision for this configuration is 85.6%, while the recall is 81.2%. The F1-score of 81.1% reinforces the model's classification consistency. Based on these results, the Chi-Square test is shown to be effective for feature selection applied to the NSL-KDD dataset.

The performance of Chi-Square does not increase linearly with the number of features. Although using 18 and 20 features produces relatively high accuracy, both remain below that of the 19-feature configuration. These findings indicate that there is a specific number of features that yield optimal performance. Excessive features may introduce redundancy that does not contribute significantly to classification. Thus, selecting an appropriate number of features becomes an important factor in improving model performance.

The data in Table 4 indicate that the ReliefF method underperforms compared to the other two feature selection techniques. ReliefF achieves its highest accuracy of 79.4% when using 20 features. In this configuration, the precision is 84.2%, and the recall reaches 79.4%. An F1-score of 79.2% shows that the improvement in performance is relatively modest. These findings suggest that ReliefF is less effective than Information Gain and Chi-Square, highlighting the importance of choosing an appropriate feature selection method to elevate the capabilities of the model.

Additionally, increasing the number of features in ReliefF results in only minor changes in model performance and overall effectiveness. The accuracy ranges from 77.9% to 79.4%, indicating that changes in feature values do not lead to significant differences. This suggests that the method is less capable of identifying features that have a substantial impact on classification performance. With such a narrow performance range, ReliefF has clear limitations. These limitations are particularly evident when the method is applied to large, complex, and real-world datasets, such as the NSL-KDD dataset.

DISCUSSIONS

The evaluation of the three boosting models in the initial phase indicates that Gradient Boosting performs best, followed by XGBoost and AdaBoost. This performance gap suggests that the sequential learning mechanism in Gradient Boosting is more effective in capturing complex patterns within the NSL-KDD dataset. Although XGBoost achieves competitive recall and F1-Scores, it still falls slightly behind Gradient Boosting in overall accuracy. In contrast, AdaBoost exhibits more pronounced weaknesses, primarily because it is susceptible to noise in the data. These findings reaffirm that Gradient Boosting is the most suitable model to serve as the foundation for subsequent optimization stages.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15650
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Feature selection was then conducted to address the overfitting issue indicated by the substantial performance divergence between training and testing. The results of feature selection using Chi-Square and Information Gain show that certain features contribute significantly to class prediction. The presence of overlapping features, such as 'dst_bytes' and 'service', across both methods underscores their strong relevance for detecting attack patterns. Differences in scoring between the two methods also reflect distinct analytical perspectives Chi-Square emphasizes correlation, whereas Information Gain focuses on entropy reduction. Consequently, both feature selection methods successfully identify influential feature subsets that improve model performance.

Applying Information Gain to the Gradient Boosting model yields consistent performance improvements across different feature counts. Although the best performance is achieved with 13 features, increasing the number of features does not always lead to further progress. This condition underscores the existence of an optimal point in determining the appropriate number of features. Moreover, a larger feature set tends to increase training and testing times, thereby affecting system efficiency. Thus, Information Gain not only improves accuracy but also provides a computationally efficient subset of features.

The Chi-Square method yields more substantial results, with accuracy reaching 81.2% when 19 features are used. These findings demonstrate that Chi-Square is highly effective at identifying categorical features that are important in the NSL-KDD dataset. However, expanding the count of features beyond the optimal threshold reduces model effectiveness due to redundancy. The increased computational burden associated with additional features also becomes a key consideration when applying this method. Overall, Chi-Square proves to be the most effective feature selection technique in this study.

The ReliefF method, which was also evaluated during the experimental stage, performed worse than the other two methods. The narrow range of accuracy across different feature configurations indicates that ReliefF is less sensitive to the complex structure of the NSL-KDD dataset. Although its stable training time is an advantage, this benefit is not accompanied by significant performance improvements. ReliefF's limited ability to identify essential features makes it less suitable for datasets with weak inter-feature relationships. Thus, this method is not suitable as a primary feature selection strategy for boosting-based intrusion detection systems.

The effectiveness of the proposed model was evaluated by comparing its performance with existing machine learning approaches with competitive performance for NIDS, considering the algorithms employed, the number of features processed, and key evaluation metrics as summarized in Table 5.

Table 5 Comparison of Model Performance with Competitive Approaches

Research	Model	Features	Accuracy	Precision	Recall	F1-Score
Proposed	GradientBoosting + Chi-Square	19	81.2%	85.6%	81.2%	81.1%
(Yuliana et al., 2023)	Decision Tree	41	80.0%	83.0%	82.0%	80.0%

The performance comparison presented in Table 5 further reinforces the contribution of the proposed approach. Compared to (Yuliana et al., 2023), which employed a Decision Tree model with 41 features, the Gradient Boosting model combined with Chi-Square feature selection demonstrates an accuracy improvement from 80.0% to 81.2%, despite using significantly fewer features, only 19. This efficiency indicates that Chi-Square effectively filters the most relevant attributes, enhancing model performance while reducing computational complexity.

Additionally, the proposed model demonstrates higher precision and F1-score values than the baseline, achieving 85.6% and 81.1%, respectively, indicating improved capability in reducing false positives while maintaining balanced detection performance. In contrast, the Decision Tree model examined in the prior study exhibits limitations in capturing non-linear relationships and complex inter-feature dependencies within the NSL-KDD dataset. Consequently, integrating Gradient Boosting with Chi-Square–based feature selection provides a more competitive and efficient alternative to traditional decision tree–based approaches. These findings align with recent advances in intrusion detection, highlighting the importance of combining ensemble learning with effective feature reduction. Beyond numerical performance comparison, it is also important to examine the practical security implications of the proposed model in real-world deployment scenarios.

The finding that feature reduction effectively reduces overfitting indicates a potential improvement in generalization capability against previously unseen or zero-day attack patterns. By preventing the model from overfitting to noisy or redundant training features, the proposed feature-efficient boosting approach enables more stable learning behavior under distributional shifts. This implication is particularly important for operational NIDS, where the ability to detect evolving and unknown attack variations is more critical than achieving high accuracy on training data alone.

From a practical security perspective, an accuracy of 81.2% combined with reduced false positives implies a more reliable alerting mechanism for network administrators, minimizing unnecessary alarms while maintaining effective detection capability. Furthermore, improved recall directly reduces the risk of false negatives, which is

*name of corresponding author

critical for preventing undetected intrusions that could compromise network security. Although the proposed model does not explicitly address zero-day adaptability, reducing feature dependency enhances robustness against unseen traffic variations by limiting reliance on redundant or noise-prone attributes. These findings indicate that feature-efficient boosting models can provide meaningful operational benefits even within classical intrusion detection frameworks. Although statistical significance testing was not conducted, the consistent performance trends across feature configurations suggest stable and reliable model behavior.

CONCLUSION

The overall findings of this study demonstrate that boosting-based ensemble algorithms, particularly Gradient Boosting combined with filter-based feature selection methods such as Chi-Square, Information Gain, and ReliefF, can significantly enhance the performance of Network Intrusion Detection Systems (NIDS). The achieved accuracy of 81.2% indicates that effective feature reduction not only improves detection capability but also reduces model complexity and mitigates overfitting, thereby improving generalization performance. Furthermore, the results confirm that selecting an appropriate subset of features plays a critical role in stabilizing boosting-based models when applied to high-dimensional intrusion detection datasets. This highlights the importance of feature efficiency in developing reliable and scalable NIDS solutions. Overall, the proposed approach demonstrates a balanced trade-off between performance and computational efficiency.

The main contributions of this study include a systematic evaluation of feature-efficient boosting models for intrusion detection on a widely used benchmark dataset, empirical evidence that Chi-Square feature selection effectively reduces overfitting while maintaining high detection performance under reduced-feature settings, and the establishment of a reproducible evaluation pipeline that ensures fair and transparent performance assessment on the NSL-KDD dataset. Despite these contributions, a key limitation of this study lies in its reliance on the NSL-KDD dataset, which may not fully reflect modern and evolving attack patterns observed in real-world network environments. As a result, the generalizability of the findings may be constrained when applied to contemporary intrusion scenarios. Therefore, future research is encouraged to incorporate more recent and diverse intrusion datasets to better capture emerging threats. Additionally, exploring advanced boosting strategies and feature selection techniques may further strengthen model robustness and applicability in dynamic, large-scale network settings.

REFERENCES

Abdullah, G. M. S., Ahmad, M., Babur, M., Badshah, M. U., Al-Mansob, R. A., Gamil, Y., & Fawad, M. (2024). Boosting-based ensemble machine learning models for predicting unconfined compressive strength of geopolymer stabilized clayey soil. *Scientific Reports*, *14*(2323). https://doi.org/10.1038/s41598-024-52825-7

Adzimi, S. N., Alfasih, H. A., Ramadhan, F. N. G., Neyman, S. N., & Setiawan, A. (2024). Implementasi Konfigurasi Firewall dan Sistem Deteksi Intrusi menggunakan Debian. *Journal of Internet and Software Engineering*, *1*(4), 12. https://doi.org/10.47134/pjise.v1i4.2681

Agustina, T., Masrizal, & Irmayanti. (2024). Performance Analysis of Random Forest Algorithm for Network Anomaly Detection using Feature Selection. *Jurnal Dan Penelitian Teknik Informatika*, *8*(2), 1116–1123. https://doi.org/10.33395/sinkron.v8i2.13625

Ahmad, T., & Aziz, M. N. (2019). Data preprocessing and feature selection for machine learning intrusion detection systems. *ICIC Express Letters*, *13*(2), 93–101. https://doi.org/10.24507/icicel.13.02.93

Ahmed, H. A., Muhammad Ali, P. J., Faeq, A. K., & Abdullah, S. M. (2022). An Investigation on Disparity Responds of Machine Learning Algorithms to Data Normalization Method. *ARO-The Scientific Journal of Koya University*, *10*(2), 29–37. https://doi.org/10.14500/aro.10970

Alsulami, B., Almalawi, A., & Fahad, A. (2022). A Review on Machine Learning Based Approaches of Network Intrusion Detection Systems. *International Journal of Current Science Research and Review*, *05*(06), 2159–2177. https://doi.org/10.47191/ijcsrr/V5-i6-47

APJII. (2024). *Data Riset Pengguna Internet di Indonesia Tahun 2024*. https://apjii.or.id/berita/d/apjii-jumlah-pengguna-internet-indonesia-tembus-221-juta-orang

Aqilah Bohani, F., Syazwani, F., Rashid, M., Mahmud, Y., & Yahya, S. R. (2024). ANALYZING THE IMPACT OF FEATURE SELECTION USING INFORMATION GAIN FOR AIRLINES' CUSTOMER SATISFACTION. *Malaysian Journal of Computing*, *9*(1), 1673–1689. https://doi.org/10.24191/mjoc.v9i1.24163

Ardana, A. (2023). Performance Analysis of XGBoost Algorithm to Determine the Most Optimal Parameters and Features in Predicting Stock Price Movement. *Jurnal Informatika Dan Teknologi Informasi*, *20*(1), 91–102. https://doi.org/10.31515/telematika.v20i1.9329

Azizah, R. A., Bachtiar, F. A., & Adinugroho, S. (2022). KLASIFIKASI KINERJA AKADEMIK SISWA MENGGUNAKAN NEIGHBOR WEIGHTED K-NEAREST NEIGHBOR DENGAN SELEKSI FITUR

*name of corresponding author

INFORMATION GAIN. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *9*(3), 605–614. https://doi.org/10.25126/jtiik.202295751

Boldini, D., Grisoni, F., Kuhn, D., Friedrich, L., & Sieber, S. A. (2023). Practical guidelines for the use of gradient boosting for molecular property prediction. *Journal of Cheminformatics*, *15*(73), 1–13. https://doi.org/10.1186/s13321-023-00743-7

Bouke, M. A., Abdullah, A., Udzir, N. I., & Samian, N. (2024). Overcoming the Challenges of Data Lack, Leakage, and Dimensionality in Intrusion Detection Systems: A Comprehensive Review. *Journal of Communication and Information Systems*, *39*(2024), 22–34. https://doi.org/10.14209/jcis.2024.3

BSSN. (2023). *LANSKAP KEAMANAN SIBER INDONESIA 2023*.

Chairunnisa, C., Ernawati Iin, & Santoni, M. M. (2022). Klasifikasi Sentimen Ulasan Pengguna Aplikasi PeduliLindungi di Google Play Menggunakan Algoritma Support Vector Machine dengan Seleksi Fitur Chi-Square. *Jurnal Informatik*, *18*(1), 69–79. https://doi.org/10.52958/iftk.v17i4.4594

Chavan, P. V., & Alone, N. V. (2025). Optimizing Intrusion Detection with Random Forest: A High-Accuracy Approach using CIC-IDS 2017. *International Journal of Computer Applications*, *187*(3), 17–22. https://doi.org/10.5120/ijca2025924816

Chhaybi, A., & Lazaar, S. (2025). Enhancing malware detection utilizing Chi-Square distribution for optimal feature selection in machine learning black box models. *Journal of Dynamics and Games*, *14*, 190–203. https://doi.org/10.3934/jdg.2025010

Dutschmann, T.-M., Kinzel, L., ter Laak, A., & Baumann, K. (2023). Large-scale evaluation of k-fold cross-validation ensembles for uncertainty estimation. *Journal of Cheminformatics*, *15*(49), 1–16. https://doi.org/10.1186/s13321-023-00709-9

Fahrezi, S. Y., Nugraha, A., Luthfiarta, A., & Primadya, N. D. (2024). Optimizing Performance of AdaBoost Algorithm through Undersampling and Hyperparameter Tuning on CICIoT 2023 Dataset. *Jurnal Ilmiah Elektroteknika*, *23*(2), 175–184. https://doi.org/10.31358/techne.v23i2.467

Fonda, H., Irawan, Y., Melyanti, R., Wahyuni, R., & Muhaimin, A. (2024). A Comprehensive Stacking Ensemble Approach for Stress Level Classification in Higher Education. *Journal of Applied Data Sciences*, *5*(4), 1701–1714. https://doi.org/10.47738/jads.v5i4.388

Freda, P. J., Ye, S., Zhang, R., Moore, J. H., & Urbanowicz, R. J. (2024). Assessing the limitations of relief-based algorithms in detecting higher-order interactions. *BioData Mining*, *17*(37), 1–18. https://doi.org/10.1186/s13040-024-00390-0

Gupta, S., Grover, D., Alzubi, A. A., Sachdeva, N., Baig, M. W., & Singla, J. (2022). Machine Learning with Dimensionality Reduction for DDoS Attack Detection. *Computers, Materials and Continua*, *72*(2), 2665–2682. https://doi.org/10.32604/cmc.2022.025048

Hakkal, S., & Lahcen, A. A. (2024). XGBoost To Enhance Learner Performance Prediction. *Computers and Education: Artificial Intelligence*, *7*, 1–10. https://doi.org/10.1016/j.caeai.2024.100254

Hussein, M. K., ALkahla, L. T., & Alqassab, A. (2024). Feature Selection Techniques in Intrusion Detection: A Comprehensive Review. *Iraqi Journal for Computers and Informatics*, *50*(1), 46–53. https://doi.org/10.25195/ijci.v50i1.462

IBM. (2025). *Cost of a Data Breach Report 2025*.

Ismanto, E., Fadlil, A., Yudhana, A., & Kitagawa, K. (2024). A Comparative Study of Improved Ensemble Learning Algorithms for Patient Severity Condition Classification. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *6*(3), 312–321. https://doi.org/10.35882/jeeemi.v6i3.452

Jaw, E., & Wang, X. (2021). Feature Selection and Ensemble-Based Intrusion Detection System: An Efficient and Comprehensive Approach. *Symmetry*, *13*(10), 1–34. https://doi.org/10.3390/sym13101764

Kaushik, B., Sharma, R., Dhama, K., Chadha, A., & Sharma, S. (2023). Performance evaluation of learning models for intrusion detection system using feature selection. *Journal of Computer Virology and Hacking Techniques*, *19*(4), 529–548. https://doi.org/https://doi.org/10.1007/s11416-022-00460-z

Malashin, I., Tynchenko, V., Gantimurov, A., Nelyub, V., & Borodulin, A. (2025). Boosting-Based Machine Learning Applications in Polymer Science: A Review. *Polymers*, *17*(4), 1–42. https://doi.org/10.3390/polym17040499

Masoodi, F., Bamhdi, A. M., & Teli, T. A. (2021). Machine Learning for Classification analysis of Intrusion Detection on NSL-KDD Dataset. *Turkish Journal of Computer and Mathematics Education*, *12*(10), 2286–2293. https://turcomat.org/index.php/turkbilmat/article/view/4768

Microsoft Digital Defense. (2024). *Microsoft Digital Defense Report 2024*.

Mulyanto, Y., Susanto, E. S., Akbar, M. I., & Idifitriani, F. (2024). Analisis Keamanan Jaringan Komputer Menggunakan Metode Intrusion Detection System (IDS) dan Firewall. *Digital Transformation Technology*, *3*(2), 864–870. https://doi.org/10.47709/digitech.v3i2.3402

Nabi, F. (2023). *Enhancing Intrusion Detection Systems: A Comparative Study of Machine Learning Techniques for Cyber Security*. https://doi.org/10.21203/rs.3.rs-3360502/v1

*name of corresponding author

Natha, S., Leghari, M., Rajput, M. A., Zia, S. S., & Shabir, J. (2022). A Systematic Review of Anomaly detection using Machine and Deep Learning Techniques. *Quaid-e-Awam University Research Journal of Engineering, Science & Technology*, *20*(1), 83–94. https://doi.org/10.52584/qrj.2001.11

Ngo, N., Michel, P., & Giorgi, R. (2024). Multivariate filter methods for feature selection with the γ-metric. *BMC Medical Research Methodology*, *24*(1), 1–22. https://doi.org/10.1186/s12874-024-02426-9

P, Poobalan., & S, Dr. P. (2022). Hybrid Sequential Feature Selection with Ensemble Boosting Class-based Classification Method. *International Journal of Recent Technology and Engineering (IJRTE)*, *11*(4), 13–18. https://doi.org/10.35940/ijrte.D7298.1111422

Putra, R. P., & Amarudin. (2025). Perbandingan Algoritma Machine Learning untuk Intrusion Detection System pada Dataset NSL-KDD. *Jurnal Sistem Informasi*, *14*(4), 1654–1664. http://sistemasi.ftik.unisi.ac.id

Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-56706-x

Regragui, Y., Mazighi, A., Ballihi, L., & Orhanou, G. (2024). Impact Evaluation of Feature Selection Algorithms on Machine Learning-Based Intrusion Detection. *Proceedings - 11th International Conference on Wireless Networks and Mobile Communications, WINCOM 2024*. https://doi.org/10.1109/WINCOM62286.2024.10656421

Saha, S., & Nandi, D. (2024). SVM-RLF-DNN: A DNN with reliefF and SVM for automatic identification of COVID from chest X-ray and CT images. *Digital Health*, *10*, 1–16. https://doi.org/10.1177/20552076241257045

Sahli, Y. (2022). A comparison of the NSL-KDD dataset and its predecessor the KDD Cup '99 dataset. *International Journal of Scientific Research and Management*, *10*(04), 832–839. https://doi.org/10.18535/ijsrm/v10i4.ec05

Saputra, N. A., Irawan, R. H., & Mahdiyah, U. (2025). Hybrid Ensemble Learning Sistem Keamanan Jaringan untuk Meningkatkan Performa Deteksi Anomali. *Jurnal Nusantara Of Engineering*, *8*(2), 361–369. https://doi.org/10.29407/noe.v8i02.25617

Schock, C., Dumler, J., & Doepper, F. (2021). Data Acquisition and Preparation - Enabling Data Analytics Projects within Production. *Procedia CIRP*, *104*, 636–640. https://doi.org/10.1016/j.procir.2021.11.107

Setiawan, Y. (2023). Data Mining berbasis Nearest Neighbor dan Seleksi Fitur untuk Deteksi Kanker Payudara. *Jurnal Pengembangan IT*, *8*(2), 89–96. https://doi.org/10.30591/jpit.v8i2.4994

Shokri, B. J., Mirzaghorbanali, A., McDougall, K., Karunasena, W., Nourizadeh, H., Entezam, S., Hosseini, S., & Aziz, N. (2024). Data-Driven Optimised XGBoost for Predicting the Performance of Axial Load Bearing Capacity of Fully Cementitious Grouted Rock Bolting Systems. *Applied Sciences (Switzerland)*, *14*(21), 1–26. https://doi.org/10.3390/app14219925

Sujon, K. M., Hassan, R. B., Towshi, Z. T., Othman, M. A., Samad, M. A., & Choi, K. (2024). When to Use Standardization and Normalization: Empirical Evidence from Machine Learning Models and XAI. *IEEE Access*, *12*, 135300–135314. https://doi.org/10.1109/ACCESS.2024.3462434

V. Priyalakshmi, & Dr. R. Devi. (2022). Evaluation of Efficient Classification Algorithm for Intrusion Detection System. *International Journal of Advanced Research in Science, Communication and Technology*, *2*(2), 39–45. https://doi.org/10.48175/ijarsct-7751

Wang, S., Balarezo, J., Kandeepan, S., Al-Hourani, A., Gomez, K., & Rubinstein, B. (2021). Machine Learning in Network Anomaly Detection: A Survey. *IEEE Access*, *4*, 1–17. https://doi.org/10.1109/ACCESS.2021.3126834

Xia, Y., Jiang, S., Meng, L., & Ju, X. (2024). XGBoost-B-GHM: An Ensemble Model with Feature Selection and GHM Loss Function Optimization for Credit Scoring. *Systems*, *12*(7), 1–26. https://doi.org/10.3390/systems12070254

Yang, F., Xu, Z., Wang, H., Sun, L., Zhai, M., & Zhang, J. (2024). A hybrid feature selection algorithm combining information gain and grouping particle swarm optimization for cancer diagnosis. *PLoS ONE*, *19*(3), 1–17. https://doi.org/10.1371/journal.pone.0290332

Yuan, Y., Shen, D., Cao, Y., Wang, X., Zhang, B., & Dong, H. (2025). An Ensemble Machine Learning Approach for High-Resolution Estimation of Groundwater Storage Anomalies. *Water (Switzerland)*, *17*(10), 1–32. https://doi.org/10.3390/w17101445

Yuliana, Supriyadi, D. H., Fahlevi, M. R., & Arisagas, M. R. (2023). Analysis of NSL-KDD for the Implementation of Machine Learning in Network Intrusion Detection System. *Journal of Informatics, Information System, Software Engineering and Applications*, *1*(1), 001–010. https://doi.org/10.20895/inista.v6i2.1389

*name of corresponding author