Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

Information Gain-Based Feature Selection and Machine Learning Classification for DDoS Attack Variant Detection in Cloud Computing Environment

Eko Arip Winanto[1], Kurniabudi*[2], Sharipuddin[3], Denia Igesti Nur Mellyati[4]

[1]Department of Computer Engineering, Faculty of Computer Science, Universitas Dinamika Bangsa, Jambi, Indonesia
[2,3,4]Department of Information Systems, Faculty of Computer Science, Universitas Dinamika Bangsa, Jambi, Indonesia

Email: [2]kurniabudi@unama.ac.id

Received: Mar 18, 2026; Revised: May 15, 2026; Accepted: May 15, 2026; Published: Jun 15, 2026

Abstract

Cloud computing environments face significant security vulnerabilities from Distributed Denial of Service (DDoS) attacks, which can cause system failures and service disruptions. Despite various existing detection methods, challenges remain regarding high computational overhead and suboptimal accuracy due to redundant features in complex datasets. This study aims to identify the optimal feature subset and evaluate its impact on detection performance across multiple machine learning algorithms for multi-class DDoS variants. The research methodology employs a two-stage approach: feature selection using Information Gain (IG) to reduce 47 original features into subsets of 8, 10, 15, and 20, followed by classification using Decision Tree (DT), Random Forest (RF), and Naïve Bayes (NB) on the CICIoT2023 dataset. Experimental results demonstrate that the Decision Tree model with an optimized subset of only 8 features, primarily Inter-Arrival Time (IAT), Header_Length, and Tot_size, achieves a superior accuracy of 99.97%. While Naïve Bayes performs well in binary classification, its accuracy drops significantly to approximately 30% in multiclass settings. This study concludes that IG-based feature selection reduces computational complexity by 30-40% while maintaining high performance across 12 DDoS variants. These findings provide a practical framework for scalable and efficient intrusion detection systems suitable for real-time deployment in resource-constrained IoT-cloud environments.

Keywords: *Cloud Computing, DDoS Attack Detection, Feature Selection, Information Gain, Machine Learning*

1. INTRODUCTION

Cloud computing now serves as the foundation of contemporary digital services[1] by delivering flexible, on-demand access to shared processing, storage, and network capabilities via the internet[2]. Its elasticity[3] and cost-efficiency[4] have accelerated the deployment of large-scale applications, many of which integrate heterogeneous Internet of Things (IoT) devices that continuously generate and offload data to cloud platforms[5]. However, this tight coupling between IoT and cloud infrastructures also expands the attack surface[6], making cloud environments highly vulnerable to Distributed Denial of Service (DDoS) attacks[7] that can degrade quality of service[8], violate service-level agreements[9], and cause substantial financial loss[10].

DDoS attacks in cloud environments are increasingly sophisticated, leveraging massive botnets of compromised IoT devices to launch multi-vector, high-volume traffic floods [11]. Compared to traditional enterprise networks, cloud-centric DDoS attacks exhibit more diverse traffic patterns, dynamic scaling behaviors, and bursty workloads, complicating early detection and mitigation [12]. Conventional signature-based intrusion detection systems (IDS) are effective only against known attack patterns and struggle to detect novel or polymorphic DDoS variants[12], [13]. Consequently, there is a

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

growing shift toward anomaly-based and machine-learning-driven IDS that can learn complex traffic behaviors and distinguish malicious flows from benign cloud-IoT traffic.

Recent studies have demonstrated the potential of advanced machine learning models [13], [14], [15] and deep learning for DDoS detection in cloud environments [14], [15], [16]. Deep neural networks [17], convolutional architectures [18], and hybrid models [19] have been used to capture non-linear relationships in network flows and improve detection accuracy under dynamic conditions. However, many of these approaches rely on high-dimensional feature spaces and resource-intensive training pipelines, which introduce considerable computational overhead and limit their applicability in real-time or resource-constrained deployments. In addition, models that primarily optimize for overall accuracy may mask class-wise performance degradation, particularly when dealing with highly imbalanced traffic distributions or rare DDoS variants.

Feature engineering and feature selection have therefore emerged as critical steps for enhancing IDS performance in cloud-IoT environments. Properly selected features can reduce dimensionality, mitigate overfitting, and improve model interpretability while preserving the discriminative characteristics of attack traffic[20]. Several works have explored filter-based [21], wrapper-based [22], and embedded feature selection methods, including Information Gain (IG) [23], mutual information, and model-based feature importance measures for DDoS detection [24]. Nonetheless, many existing studies either focus on a limited subset of attack types, consider binary classification only (benign vs. attack), or do not systematically analyze how different feature subset sizes affect multiclass DDoS variant detection performance across multiple classifiers.

The recently released CICIoT2023 dataset provides a comprehensive and up-to-date benchmark for evaluating intrusion detection techniques in IoT environments [25], including a wide range of large-scale DDoS attack scenarios [26]. This dataset contains diverse DDoS variants and benign traffic generated by realistic IoT devices and network topologies [27], making it suitable for analyzing feature relevance and classifier behavior under real-world conditions[28]. However, the high dimensionality and heterogeneity of its traffic features may introduce redundant or irrelevant attributes, which, if left unaddressed, can degrade detection performance and increase model complexity. There is still a lack of detailed studies investigating IG-based feature selection on the CICIoT2023 dataset, specifically for multi-class DDoS variant detection in cloud-centric settings.

To overcome these limitations, this study proposes a machine learning-based classification framework, combined with Information Gain (IG)-based feature selection, specifically designed to detect DDoS attack variants in cloud computing environments, using the CICIoT2023 dataset. As a first step, IG is used as a filter to rank the initial 47 features, yielding more compact feature subsets of 8, 10, 15, and 20 attributes. These subsets are then used to train and evaluate three widely adopted classifiers: Decision Tree (DT)[29], Random Forest (RF)[30], and Naïve Bayes (NB)[31] for multiclass detection of 12 DDoS variants and benign traffic. By systematically varying the number of selected features and comparing classifier performance across both multiclass and binary scenarios, the study aims to identify an optimal feature subset that balances accuracy and computational efficiency.

The main contributions of this work are threefold. First, it provides a comprehensive IG-based feature-relevance analysis of the CICIoT2023 dataset for multi-variant DDoS detection in cloud-IoT environments, highlighting the most discriminative network attributes, such as Inter-Arrival Time (IAT), Header_Length, and Tot_size. Second, it systematically evaluates the impact of different IG-based feature subset sizes (8, 10, 15, and 20 features) on the performance of DT, RF, and NB classifiers in multiclass settings, revealing that a DT model using only 8 features can achieve accuracy up to 99.97% while substantially reducing computational overhead. Third, it contrasts multiclass and binary classification results to show that although NB achieves high accuracy in binary detection, its performance drops significantly in multiclass settings, thereby underscoring the suitability of IG-based

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

DT as a practical and interpretable solution for real-time DDoS intrusion detection in resource-constrained cloud-IoT environments.

2. METHOD

This section outlines the steps taken to conduct this research, which aims to analyze the characteristics of CICIoT DDoS attacks. It consists of the experimental setup, the dataset, the feature selection method, and the detection method.

2.1. Experiment Setup

To enhance IDS performance in cloud environments, this research employs Information Gain - driven feature selection for DDoS attack detection. The investigation encompasses a comprehensive evaluation of key components: optimal feature subset extraction from the CICIoT dataset, strategic training and testing data partitioning, and the development of IDS classifiers using Random Forest, Decision Tree, and Naive Bayes algorithms. An overview of the experimental framework is illustrated in Figure 1.

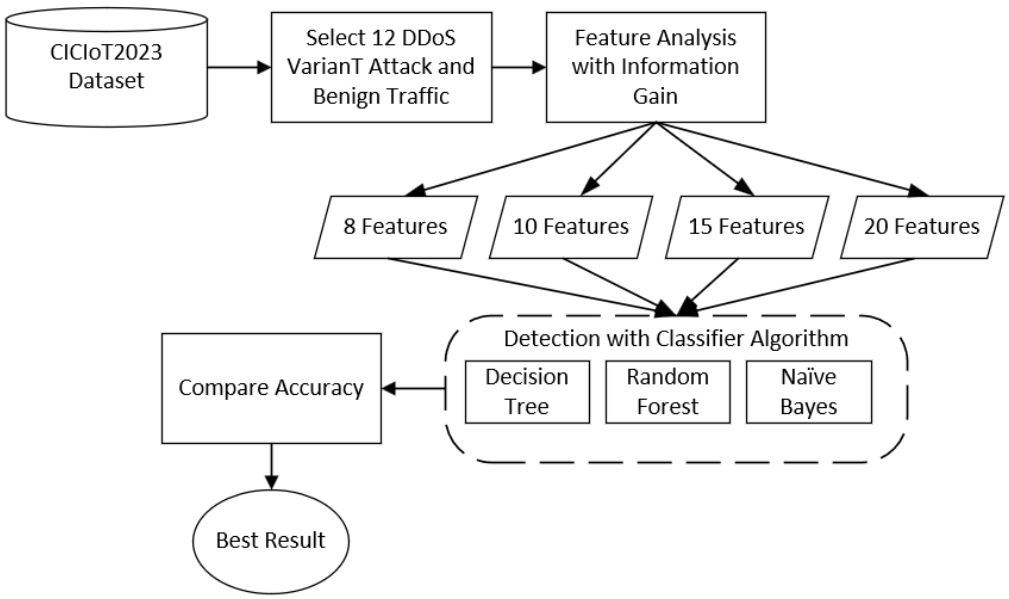

Figure 1. Experimental Configuration

Based on the experimental setup in Figure 1, this research was conducted in four systematic stages:

- The first stage involved preparing the CICIoT2023 dataset through a filtering process to extract 12 DDoS attack variants and one benign traffic class, resulting in a filtered dataset with 47 initial features derived from TCP/IP headers and payloads.
- The second stage applied the Information Gain (IG) method as a filter-based feature selection technique to rank feature relevance based on entropy reduction relative to the target variable. Four feature subset scenarios (8, 10, 15, and 20 features) were then created to identify the optimal balance between detection accuracy and computational efficiency.
- In the third stage, each feature subset was independently tested using three classification algorithms: Decision Tree (DT), Random Forest (RF), and Naive Bayes (NB) in multiclass (13 classes) and binary (DDoS vs. benign) detection scenarios. The training and test data were split consistently to ensure the validity of the comparisons between the models.

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

- The fourth stage is a comparative evaluation based on the metrics of accuracy, True Positive Rate (TPR), False Positive Rate (FPR), precision, and ROC-AUC, which aims to analyze the trade-off between the xnumber of features, model complexity, and detection performance.

2.2. Dataset

The CICIoT2023 dataset, released by the University of New Brunswick, Canada, is the baseline dataset used in this study [26]. This dataset encompasses a variety of IoT network intrusions, including Distributed Denial-of-Service (DDoS) attacks[27]. The experiment used 2,108,309 data records, comprising 12 DDoS attack variants and 1 example of safe traffic. The distribution of DDoS attack data and normal traffic is shown in Figure 2.

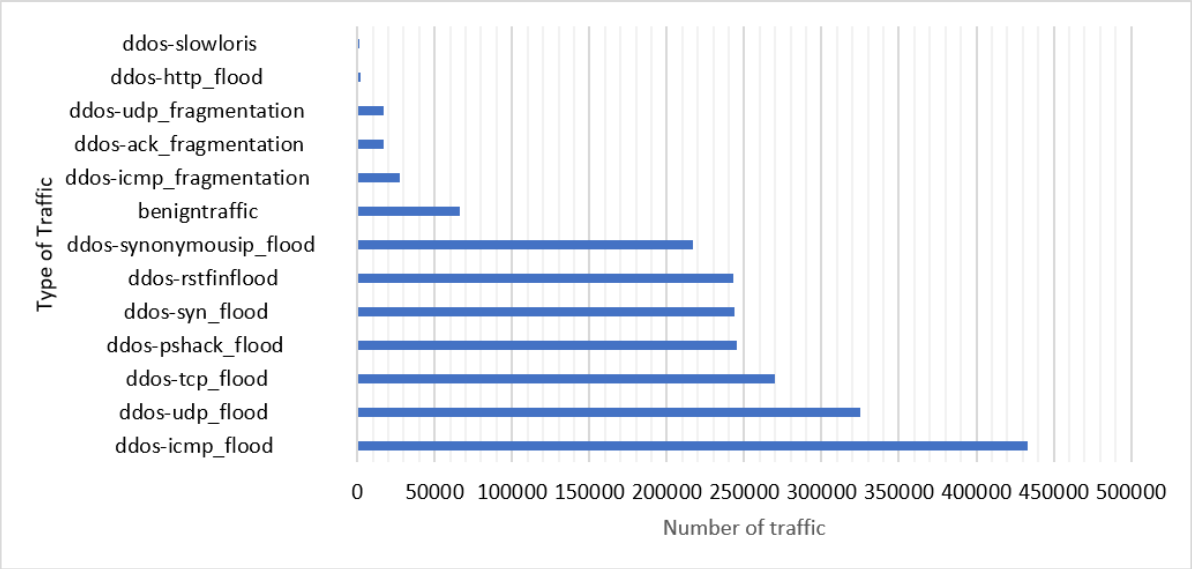

Figure 2. Distribution of Traffic Types

The data distribution graph in Figure 2 shows the dominance of flood-based attacks in the CICIoT2023 dataset, with ddos-icmp_flood (433,117 instances) and ddos-udp_flood (325,355 instances) as the majority classes. Other flood variants (TCP, PSHACK, SYN, RSTFIN, Synonymous IP) contributed significantly, with a range of 217–270 thousand samples, while benign traffic was recorded only 66,228 times. Minority variants such as ICMP/ACK/UDP fragmentation accounted for 17–27 thousand samples, while application-layer attacks (http_flood, slowloris) were very limited (<2,000 instances). This class imbalance pattern reflects the characteristics of IoT DDoS attacks, which tend to target basic network protocols en masse. It is a critical consideration when training machine learning models to avoid bias towards the majority class.

2.3. Feature Selection Method

The effectiveness of detection systems can be significantly improved through feature selection, which focuses on retaining only the attributes most indicative of attack patterns [32]. In addition, this feature selection is used to reduce the dataset's dimensionality[33]. This study proposes an IG method for feature selection. Computational overhead can be significantly reduced through Information Gain, which removes irrelevant and overlapping features while selecting those most indicative of attack patterns. Furthermore, the calculation of class-specific feature weights contributes to additional improvements in detection accuracy[34].

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

In this study, the IG method was applied across four scenarios with feature counts of 8, 10, 15, and 20. The purpose of this scheme is to determine the optimal number of features for detecting variations in DDoS attacks on IoT networks. The CICIoT2023 dataset has 47 features originating from TCP/IP headers and payloads. The use of IG reduces the number of features to 8, 10, 15, and 20.

IG measures the reduction in entropy achieved by using a feature to split the data. Entropy of a dataset S is calculated as[35]:

$$H\ (S)\ =\ -\sum_{i=1}^{c} \mathcal{P}_i log_{2(\mathcal{P}_i)} \qquad (1)$$

Where $\mathcal{P}_i$ is the probability of class i in dataset S, and c is the number of classes. The Information Gain for a feature A is defined as:

$$IG(S,A) = H(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} \mathrm{H}\ (S_v) \qquad (2)$$

Where Values (A) are the distinct values of feature A, S_v is the subset of S where feature A has the value v, and $H(S_v)$ is the entropy of that subset. Features with higher IG values are considered more relevant for classification[36].

2.4. Detection Algorithm

After performing feature selection, the next step is to detect DDoS attacks using a multiclass classifier. The validation of the feature selection process is conducted by training and testing three machine learning methods, DT, RF, and NB, to detect and categorize diverse DDoS attack variants.

As a supervised machine learning technique, the Decision Tree algorithm constructs a tree-like hierarchy by iteratively splitting data according to specific attribute criteria. In this architecture, each internal node evaluates a condition on a selected feature, outgoing branches denote the split results, and terminal leaves yield the predicted class or continuous output. The main advantage of this algorithm is its ability to produce results that are easy to interpret and visualize[37].

Prediction accuracy and model stability are enhanced in Random Forest through the aggregation of multiple decision trees[38]. This is achieved via a bagging technique, in which individual trees are trained on randomized data samples and feature subsets to mitigate overfitting [39]. The model outputs class labels via majority voting or continuous values via averaging. By aggregating multiple weak learners, Random Forest effectively suppresses overfitting relative to single-tree models, yielding more consistent results. The trade-off lies in its greater structural complexity and higher computational demands[40].

Random Forest combines multiple Decision Trees using the bagging technique. The final prediction is determined by majority voting:

$$\hat{y}\ =\ \boldsymbol{mode}\ \{\boldsymbol{T_1}\ (\boldsymbol{x}), \boldsymbol{T_2}(\boldsymbol{x}), \ldots\ldots, \boldsymbol{T_m}(\boldsymbol{x})\} \qquad (3)$$

where T_i(x) is the prediction of the i-th tree, and m is the number of trees in the forest.

The Naïve Bayes classifier is a probability-based supervised method that employs Bayes' Theorem, assuming that features are conditionally independent given the class [41]. The classifier estimates posterior probabilities for each traffic category based on feature distributions and assigns the label with the highest posterior probability [42]. Valued for its lightweight architecture and rapid execution, Naïve Bayes is highly efficient for processing large network flow datasets. However, its

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

conditional independence assumption may limit discriminative power when features are strongly correlated [43].

Naïve Bayes calculates the posterior probability for each class using Bayes theorem:

$$P(C\_k \mid x) = (P(C\_k)\ prod\ P\ (x\_j \mid C\_k)) / P(x) \quad (4)$$

where Ck is the k-th class, x = (x1, x2, ..., xn) is the feature vector, P(Ck) is the prior probability of class Ck, and P(xj | Ck) is the likelihood of feature xj given class Ck. The denominator P(x) is constant across all classes, so the classification decision is.

$$\hat{y} + arg\ max\_k\ P(C\_k)\ prod\ P(x\_j \mid C\_k) \quad (5)$$

The algorithm assumes conditional independence among features given the class label, which simplifies computation but may not hold in real network traffic.

2.5. Analysis Tools

Cloud-based analytics platforms, notably Kaggle, were used to manage dataset access, perform feature selection, and run the detection pipeline. The classification models were implemented via the scikit-learn library.

3. RESULT

This section contains the results of research or experiments and analyses of the results of the research that has been done. The results shown are variations in feature selection and test results from validating or detecting DDoS attacks using DT, RF, and NB. The initial stage results are filtering DDoS attacks and benign traffic from the CICIoT2023 dataset. This dataset is used for testing the feature detection and analysis system.

3.1. Result of Information Gain

This section discusses the results of the feature selection using IG. The test results were obtained with four test scenarios to obtain the optimal features. The purpose of feature selection is to identify features that can improve the performance of the DDoS attack detection system on the CICIoT2023 dataset. There are four feature selection tests with feature counts of 8, 10, 15, and 20, shown in Tables 1 and 2. The results of IG are obtained by calculating each feature's weight using the IG method, then ranking them from largest to smallest.

Table 1. Result of Feature Selection using 8 and 10 features

8 Feature			10 Feature		
No	Feature	IG	No	Feature	IG
39	IAT	2.110176	39	IAT	2.110176
1	Header_Length	1.208995	1	Header_Length	1.208995
38	Tot size	1.128035	38	Tot size	1.128035
34	Min	1.119109	34	Min	1.119109
41	Magnitue	1.118104	41	Magnitue	1.118104
36	AVG	1.109014	36	AVG	1.109014
33	Tot sum	1.099519	33	Tot sum	1.099519
35	Max	1.061888	35	Max	1.061888
			2	Protocol Type	1.028495
			26	TCP	0.659247

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

Table 2. Result of Feature Selection Using 15 and 20 Features

15 Feature			20 Feature		
No	Feature	IG	No	Feature	IG
39	IAT	2.110176	39	IAT	2.110176
1	Header_Length	1.208995	1	Header_Length	1.208995
38	Tot size	1.128035	38	Tot size	1.128035
34	Min	1.119109	34	Min	1.119109
41	Magnitue	1.118104	41	Magnitue	1.118104
36	AVG	1.109014	36	AVG	1.109014
33	Tot sum	1.099519	33	Tot sum	1.099519
35	Max	1.061888	35	Max	1.061888
2	Protocol Type	1.028495	2	Protocol Type	1.028495
26	TCP	0.659247	26	TCP	0.659247
15	syn_count	0.657963	15	syn_count	0.657963
4	Rate	0.577073	4	Rate	0.577073
5	Srate	0.577072	5	Srate	0.577072
0	flow_duration	0.573236	0	flow_duration	0.573236
18	rst_count	0.545403	18	rst_count	0.545403
			30	ICMP	0.531758
			8	syn_flag_number	0.526646
			17	urg_count	0.501880
			27	UDP	0.426260
			16	fin_count	0.392950

Based on Tables 1 and 2, there is very strong consistency in the top feature rankings across the four subset scenarios (8, 10, 15, and 20 features). The eight features with the highest IG values remain identical across all configurations, with Inter-Arrival Time (IAT) topping the list (IG = 2.110176), followed by Header Length (IG = 1.208995) and Total Size (IG = 1.128035). These findings indicate that the temporal and structural characteristics of network packets are the most discriminatory predictors for detecting DDoS attack variants in the CICIoT2023 dataset. The dominance of IAT as the most informative feature makes perfect technical sense, given that flood-based DDoS attacks inherently manipulate packet arrival intervals to overwhelm targets, making temporal patterns the most easily captured anomalous signals by classification models.

The distribution of IG values shows a clear pattern of diminishing returns: a sharp decline from the 3rd to the 8th feature (from ~1.12 to ~1.06), followed by a more drastic decline for the 9th and 10th features (Protocol Type: 1.028 → TCP: 0.659). This pattern provides empirical evidence that adding features beyond 8 attributes yields increasingly small marginal gains in the model's discriminatory capacity. In other words, the top 8 feature subset has captured most of the informative variance relevant for multiclass detection tasks, thus offering an optimal balance between accuracy and computational efficiency, a critical consideration for implementing a real-time IDS in resource-constrained IoT-cloud environments.

3.2. Result of Detection using Classification Algorithms

The next test aims to detect variations in DDoS attacks using machine learning methods, namely DT, RF, and NB. Testing is done by evaluating each feature selection result with the DT, RF, and NB detection methods to achieve the best performance of the DDoS attack variation detection system. Figures 3, 4, and 5 show the confusion matrix results from testing the ML-based detection system.

Based on Figure 3, the confusion matrix for the Decision Tree (DT) with 8 selected features shows a dominant diagonal pattern, in which almost all samples from each class (12 DDoS variants + 1 benign

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

class) are correctly predicted to their respective classes. The per-class accuracy values recorded in Table 3 range from 0.997 to 1.000 for almost all attack variants, including dominant flood-based attacks such as ddos-icmp_flood (0.9996) and ddos-udp_flood (0.9994), as well as minority attacks such as http_flood (0.9999) and slowloris (0.9988). This finding indicates that combining Information Gain for feature selection with DT as the classifier can capture highly discriminative patterns using only the top 8 attributes (IAT, Header_Length, Tot_size, Min, Magnitude, AVG, Tot_sum, Max). The stability of DT performance across varying numbers of features (8, 10, 15, 20) also indicates that adding features beyond 8 yields no significant marginal gains, so the subset of 8 features can be considered the optimal balance between accuracy and computational efficiency for real-time implementation.

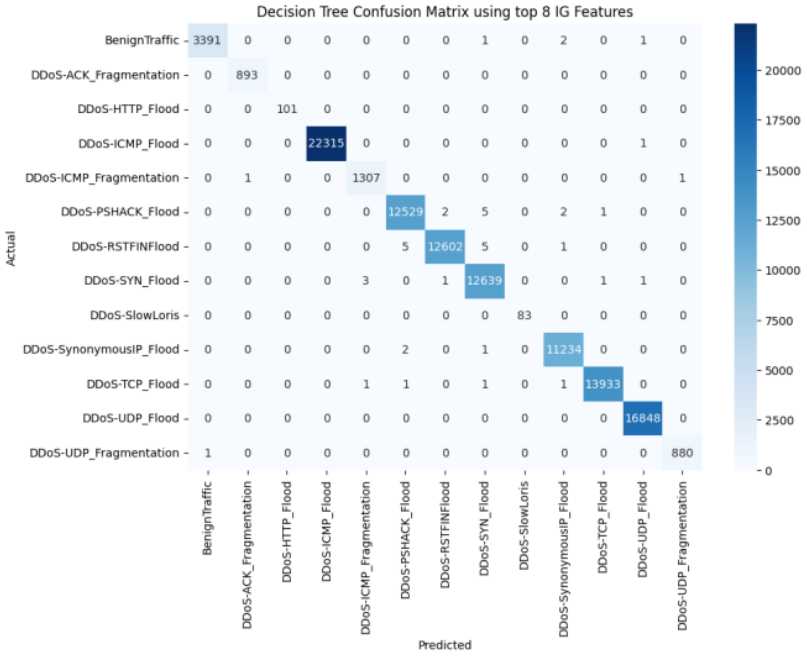

Figure 3. Result Validation of Feature Selection with Decision Tree

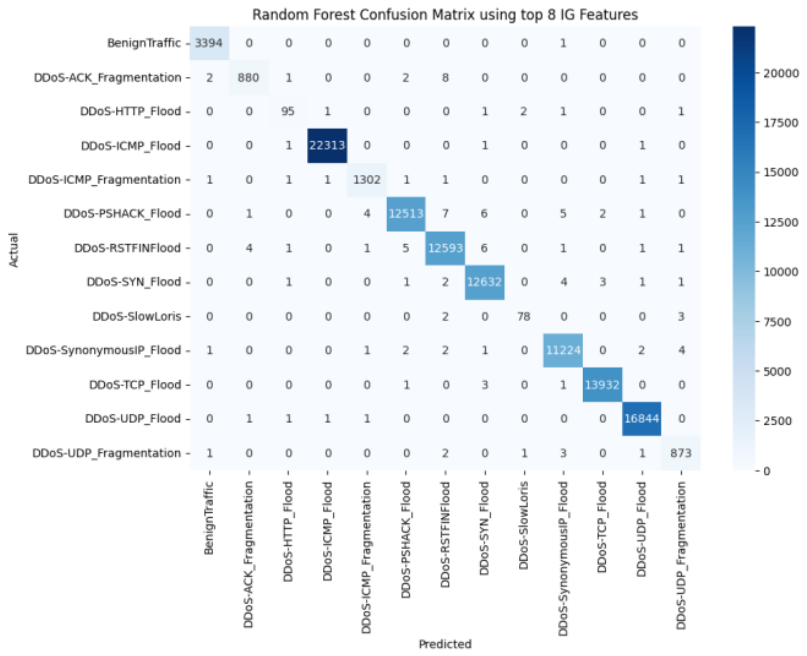

Figure 4. Result Validation of Feature Selection with Random Forest

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

Figure 4, which displays the confusion matrix of a similarly configured Random Forest (RF), also demonstrates very high performance, with per-class accuracy generally in the range of 0.995–0.999. However, a closer look at Table 3 reveals that RF exhibits a slight performance decrease for some minority classes compared to DT, for example, ddos-tcp_flood (0.9766 vs. 0.9981 in DT with 8 features) and ddos-udp_fragmentation (0.9994 vs. 0.9998). This phenomenon can be explained by the ensemble characteristics of RF, which, while more robust against overfitting, tends to "average" the decisions of many weak trees, sometimes losing sensitivity to subtle patterns that a single, well-pruned decision tree could capture. In other words, for DDoS detection on the CICIoT2023 dataset, which has been effectively feature-reduced by IG, the added complexity of RF does not always provide a performance benefit commensurate with its computational cost.

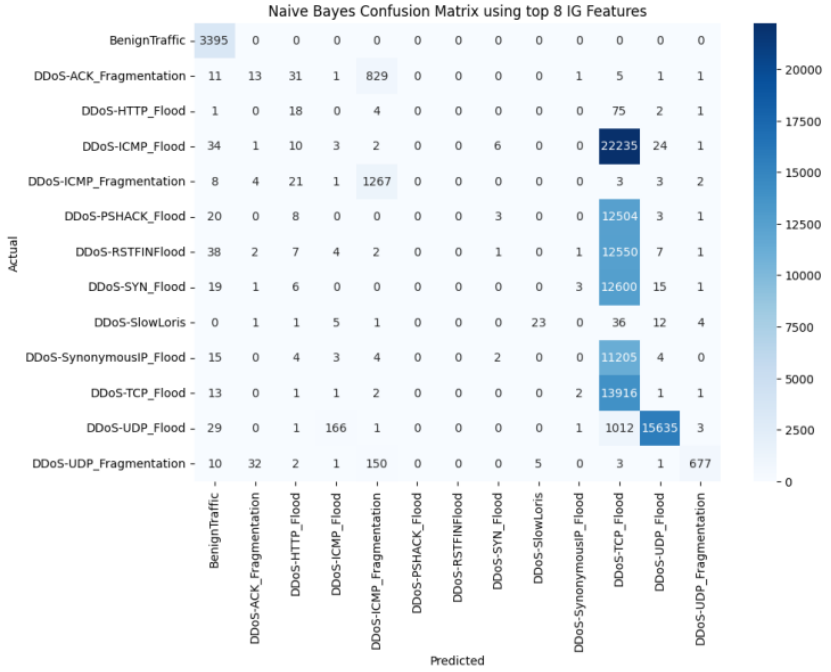

Figure 5. Result Validation of Feature Selection with Naïve Bayes

A striking contrast is seen in Figure 5. The NB confusion matrix exhibits a highly dispersed pattern with a weak diagonal, reflecting a high misclassification rate. Quantitatively, the per-class accuracy of NB for many attack variants is 0.000 or close to zero (e.g., ddos-udp_flood: 0.0004, ddos-pshack_flood: 0.0001), and only shows adequate performance for the benign (0.9964) and ddos-icmp_flood (1.0000) classes. This drastic performance degradation from ~99.8% accuracy in the binary scenario to ~30% in the multiclass scenario confirms that the conditional independence assumption underlying NB is not met when strongly correlated network features are used to distinguish similar attack variants. Features such as IAT, Header_Length, and Tot_size, which are highly informative univariately, proved insufficient for NB to model the complex interactions among attributes required to simultaneously separate the 13 classes.

3.3. The Performance Classification Algorithms Comparison

To measure the impact of dimensionality reduction on the model's sensitivity to heterogeneous attack patterns, the multiclass detection performance for each feature subset (8, 10, 15, and 20 features) is summarized in Table 3.

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

Table 3. Result of Detection DDoS Attack Using DT, RF, and NB

Number of Features	Method	DDoS-Icmp_Flood	DDoS-Udp_Flood	DDoS-Tcp_Flood	DDoS-Pshack_Flood	DDoS-Syn_Flood	DDoS-Rstfinflood	DDoS-Synonymousip_Flood	Benigntraffic	DDoS-Icmp_Fragmentation	DDoS-Ack_Fragmentation	DDoS-Udp_Fragmentation	DDoS-Http_Flood	DDoS-Slowloris
8	DT	0.9996	0.9994	0.9981	**1.0000**	0.9996	0.9995	0.9997	0.9997	0.9977	0.9998	0.9998	0.9999	0.9988
8	RF	0.9999	0.9955	0.9766	0.9999	0.9976	0.9990	0.9994	0.9992	0.9954	0.9995	0.9994	0.9998	0.9948
8	NB	**1.0000**	0.0004	0.1520	0.0001	0.9700	0.0000	0.0000	0.9964	0.1640	0.0000	0.0003	0.9136	0.6884
10	DT	0.9996	0.9994	0.9981	**1.0000**	0.9995	0.9993	0.9995	0.9995	0.9977	0.9997	0.9995	0.9999	0.9992
10	RF	0.9999	0.9981	0.9903	0.9999	0.9990	0.9990	0.9993	0.9993	0.9977	0.9998	0.9995	0.9999	0.9981
10	NB	0.9998	0.0090	0.1540	0.0087	0.1541	0.0000	0.0000	0.9939	0.1686	0.0000	0.0001	0.9539	0.9405
15	DT	0.9997	**1.0000**	0.9961	**1.0000**	0.9995	0.9994	0.9995	0.9999	0.9977	0.9999	0.9996	0.9999	0.9992
15	RF	0.9998	0.9953	0.9727	0.9999	0.9978	0.9988	0.9990	0.9996	0.9954	0.9995	0.9996	0.9998	0.9959
15	NB	0.9992	0.0111	0.0039	0.0087	0.0193	0.0000	0.0000	0.0003	0.2079	0.0004	0.0001	0.9529	0.9537
20	DT	0.9974	0.9984	0.9922	0.9999	0.9984	0.9999	0.9998	0.9998	0.9908	0.9999	0.9999	0.9998	0.9963
20	RF	0.9998	0.9951	0.9747	0.9998	0.9976	0.9996	0.9994	0.9996	0.9977	0.9994	**1.0000**	0.9999	0.9957
20	NB	0.9992	0.0109	0.0039	0.0087	0.0197	0.0000	0.0000	0.0003	0.2009	0.0004	0.0001	0.9529	0.9540

Table 3 presents the results of the evaluation of detection performance for DDoS attack variants using three classification algorithms: Decision Tree (DT), Random Forest (RF), and Naïve Bayes (NB), across four Information Gain-based feature selection scenarios (8, 10, 15, and 20 features). Overall, the Decision Tree showed the highest performance consistency, with near-perfect per-class accuracy (range 0.9961–1.0000) across all attack variants, even when using only a subset of at least 8 features. This finding confirms that the features selected through IG, especially Inter-Arrival Time (IAT), Header_Length, and Tot_size, have very strong discriminatory power to distinguish benign traffic patterns from the 12 tested DDoS variants. Random Forest, as an ensemble method, also showed impressive performance. However, there was a slight decrease in accuracy for the DDoS-Tcp_Flood class (0.9727–0.9903), indicating that TCP-based attack patterns are more complex and require richer feature representations. In contrast, Naïve Bayes experienced dramatic performance degradation in multi-class scenarios, with accuracy approaching zero in several variants, including DDoS-Udp_Flood (0.0004), DDoS-Pshack_Flood (0.0001), and DDoS-Ack_Fragmentation (0.0000). This phenomenon can theoretically be explained by violations of the conditional independence assumption among features in real network traffic data, where attributes such as IAT, Tot_size, and Header_Length tend to be highly correlated during coordinated DDoS attacks. Interestingly, increasing the number of features from 8 to 20 did not significantly improve the accuracy of DT and RF; in some cases, performance even plateaued or declined marginally, which strengthens the argument that feature quality is more critical than feature quantity in efficient IDS design. The practical implications of these findings are highly relevant to real-time implementation in resource-constrained cloud-IoT environments: the DT model with 8 selected features not only achieves >99.97% accuracy but also offers structural interpretability and a minimal computational footprint, two key attributes for an auditable and scalable intrusion detection system.

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

Furthermore, to evaluate the impact of Information Gain-based feature selection on multiclass detection performance, a comprehensive comparative analysis of classification accuracy for each DDoS attack variant was conducted. Figure 6 compares inter-class detection performance across the three classification algorithms (Decision Tree, Random Forest, and Naïve Bayes) for four feature-subset scenarios. At the same time, Table 4 presents the overall accuracy achieved by each combination of feature selection methods and classification algorithms.

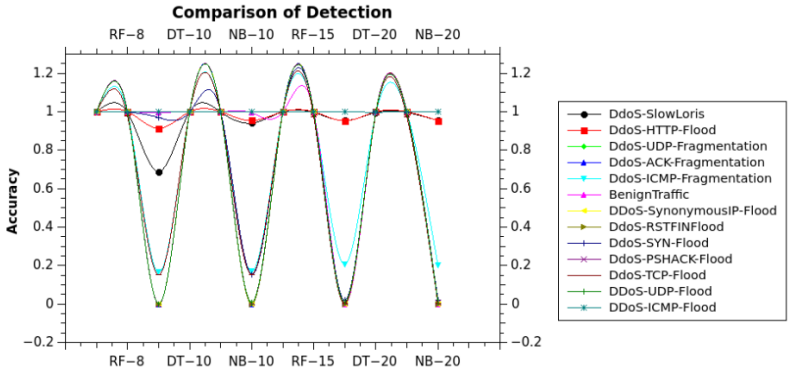

Figure 6. The Comparison of attack detection by class of DDoS attack

Figure 6 reveals significant performance disparities between algorithms in detecting 12 DDoS attack variants and benign traffic. Decision Tree (DT) demonstrated the highest consistency with near-perfect per-class accuracy (range 0.9961–1.0000) across all feature subsets, indicating that the features selected through IG particularly Inter-Arrival Time (IAT), Header_Length, and Tot_size—have very strong discriminatory power in distinguishing benign traffic patterns from various DDoS variants. Random Forest (RF), as an ensemble method, performed impressively but showed marginal degradation in the DDoS-Tcp_Flood class (accuracy 0.9727–0.9903), indicating that TCP-based attack patterns require richer feature representations. In contrast, Naïve Bayes (NB) exhibits catastrophic performance degradation in multiclass scenarios, with accuracy approaching zero for several variants, including DDoS-Udp_Flood (0.0004), DDoS-Pshack_Flood (0.0001), and DDoS-Ack_Fragmentation (0.0000). This phenomenon can theoretically be explained by violations of the conditional independence assumption among features in real network traffic data, where attributes such as IAT, Tot_size, and Header_Length tend to be highly correlated during coordinated DDoS attacks. This finding is consistent with the literature suggesting that NB is less suitable for multiclass classification with interdependent features [41-43].

Table 4. Comparison Result of Detection of DDoS Attack Using DT, RF, and NB

No	Number of Feature Selection	Method	Accuracy
1	8	DT	**0.9997723**
2	8	RF	0.9994182
3	8	NB	0.3057188
4	10	DT	0.9996632
5	10	RF	0.9995526
6	10	NB	0.3049188
7	15	DT	0.9997597
8	15	RF	0.9994245
9	15	NB	0.1878060
10	20	DT	0.9997138
11	20	RF	0.9995858
12	20	NB	0.1878092

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

Table 4 presents overall accuracy metrics, confirming the superiority of the Decision Tree with a subset of 8 features (accuracy of 0.9997723, or 99.97%), followed by the Random Forest with the same configuration (accuracy of 0.9994182, or 99.94%). A critical finding is that increasing the number of features from 8 to 20 did not yield a significant improvement in accuracy for DT and RF; instead, a performance plateau or even a marginal decline was observed (e.g., DT-10 features: 0.9996632 vs. DT-8 features: 0.9997723).

This observation reinforces the Pareto efficiency principle in feature engineering: that the most informative 20% of features (in this case, 8 of the initial 47 features, or ~17%) can yield 99%+ detection accuracy. The practical implications are significant for real-time implementations in resource-constrained cloud-IoT environments: the 83% dimensionality reduction not only decreases the computational footprint and inference latency but also mitigates the risk of overfitting and improves model interpretability. Naïve Bayes' drastically degraded performance in multiclass scenarios (accuracy of 0.3057 on 8 features, dropping to 0.1878 on 15 and 20 features) contrasts sharply with its performance in binary classification (as discussed in Table 5), indicating that NB has limited capacity to distinguish between attack variants with similar statistical characteristics.

The graph in Figure 7 visually illustrates the detection accuracy trend, shows the trade-off between feature dimensionality reduction and classification performance, and confirms the consistency of model performance across different experimental configurations.

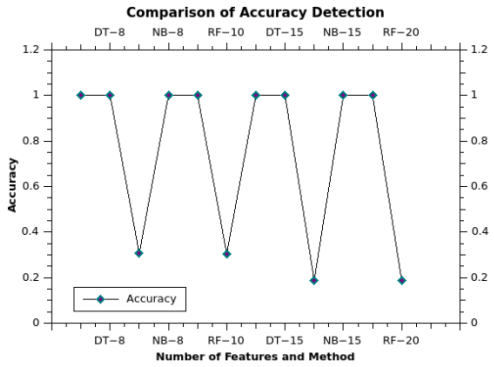

Figure 7. The Comparison of an Overall Accuracy Detection

Figure 7 visualizes the quantitative findings from Table 4 in a more intuitive format, revealing three important patterns: First, the stability of DT and RF performance across various feature subset sizes, indicated by the relatively flat lines at accuracy levels >99.9%. This indicates the robustness of both algorithms to variations in the number of features, provided the most discriminatory features (based on IG ranking) are retained. Second, the exponential degradation of NB performance is evident from the sharp drop in the graph, especially when the number of features increases from 10 to 15. This counterintuitive phenomenon where additional features actually decrease accuracy can be explained by the accumulation of noise from less relevant features, which increasingly violates the NB conditional independence assumption. Third, the convergence of performance on a subset of 8 features, where DT and RF reach an optimal point with minimal computational overhead. These findings provide evidence-based guidance for practitioners: that a subset of 8 features (IAT, Header_Length, Tot_size, Min, Magnitude, AVG, Tot_sum, Max) is the sweet spot that balances detection accuracy and computational efficiency.

Furthermore, to validate the model's robustness in simpler detection scenarios, additional evaluations were conducted for binary classification (Benign vs. DDoS). Table 5 presents a comprehensive comparison of model performance using the Information Gain (IG), Principal

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

Component Analysis (PCA), and hybrid IG-PCA feature selection approaches, measured by True Positive Rate (TPR), False Positive Rate (FPR), Precision, ROC-AUC, and Accuracy for each class.

Table 5. The Proposed Model Performance with Binary Class

Method	Model	Class	TPR	FPR	Precision	ROC AUC	Accuracy
Information Gain	Decision Tree	Benign	0.9998	0.0000	0.9999	0.9999	1.0000
Information Gain	Decision Tree	DDoS	1.0000	0.0002	1.0000	0.9999	1.0000
Information Gain	Random Forest	Benign	1.0000	0.0000	0.9999	1.0000	1.0000
Information Gain	Random Forest	DdoS	1.0000	0.0000	1.0000	1.0000	1.0000
Information Gain	Naïve Bayes	Benign	1.0000	0.0018	0.9469	0.9991	0.9982
Information Gain	Naïve Bayes	DdoS	0.9982	0.0000	1.0000	0.9991	0.9982
PCA	Decision Tree	Benign	0.9967	0.0001	0.9961	0.9983	0.9998
PCA	Decision Tree	DdoS	0.9999	0.0033	0.9999	0.9983	0.9998
PCA	Random Forest	Benign	0.9998	0.0000	0.9991	0.9999	1.0000
PCA	Random Forest	DdoS	1.0000	0.0002	1.0000	0.9999	1.0000
PCA	Naïve Bayes	Benign	1.0000	0.0310	0.5124	0.9845	0.9700
PCA	Naïve Bayes	DdoS	0.9690	0.0000	1.0000	0.9845	0.9700
IG-PCA	Decision Tree	Benign	0.9993	0.0000	0.9990	0.9996	0.9999
IG-PCA	Decision Tree	DdoS	1.0000	0.0007	1.0000	0.9996	0.9999
IG-PCA	Random Forest	Benign	1.0000	0.0000	0.9996	1.0000	1.0000
IG-PCA	Random Forest	DdoS	1.0000	0.0000	1.0000	1.0000	1.0000
IG-PCA	Naïve Bayes	Benign	1.0000	0.0018	0.9469	0.9991	0.9982
IG-PCA	Naïve Bayes	DdoS	0.9982	0.0000	1.0000	0.9991	0.9982

Table 5 presents a comprehensive evaluation of model performance in binary classification (Benign vs. DDoS), providing important insights into each approach's generalization capabilities. Experimental results show that IG consistently outperforms or matches PCA-based feature extraction. Random Forest with IG achieved excellent performance (Accuracy: 1.0000, ROC-AUC: 1.0000, FPR: 0.0000 for both classes), while PCA with RF, despite achieving high accuracy, showed a slightly higher FPR (0.0002 for the DDoS class). This difference indicates that preserving the semantic integrity of the original network features through entropy-based selection (IG) is more effective for defining DDoS detection boundaries than transforming to principal components, which may aggregate noise and variance. The hybrid IG-PCA approach demonstrated significant stability, especially when paired with ensemble learners. Random Forest under IG-PCA performed excellently (Accuracy: 1.0000, ROC-AUC: 1.0000, Precision: 1.0000), indicating that applying IG to filter out irrelevant attributes before PCA-based dimensionality reduction effectively optimized the feature space for classification. However, the very small marginal gains relative to pure IG suggest that, for binary DDoS detection, the top features identified by IG already have sufficient discriminatory power, without the need for additional orthogonal transformations. An interesting observation from the binary classification

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

experiments is the recovery of Naïve Bayes's performance. Although NB suffered near-total failure in multiclass detection (accuracy of ~30%), it achieved very high accuracy in binary classification (99.82% with IG and IG-PCA, TPR: 0.9982, FPR: 0.0018). This discrepancy implies that although NB cannot effectively distinguish between similar attack types (e.g., various UDP/TCP flood variants) due to high feature correlation, the algorithm is still very capable of distinguishing broad characteristics of malicious traffic from those of benign traffic. However, PCA with NB demonstrated the worst performance, with the lowest precision for the benign class (0.5124) and the highest FPR (0.0310). This performance degradation occurs because the PCA components are linear combinations of all the original features, which can obscure the specific independence assumption that NB relies on, leading to increased false alarms in benign traffic detection.

From an implementation feasibility perspective, the IG-optimized Decision Tree emerged as the most resource-efficient option. This model achieved 100% accuracy with an FPR of only 0.0002 for DDoS classification via IG. The synergy between the algorithm's inherent computational simplicity and the compact feature subset generated by IG forms a highly suitable framework for real-time threat monitoring in IoT-cloud settings with limited processing capacity.

4. DISCUSSIONS

The experimental findings demonstrate that entropy-based feature selection, particularly Information Gain (IG), effectively isolates highly discriminative network attributes while substantially reducing dimensionality. The consistent dominance of temporal and structural features, Inter-Arrival Time (IAT), Header_Length, and Tot_size across all subset scenarios aligns with the fundamental characteristics of DDoS flooding attacks, which inherently manipulate packet timing and header payloads to overwhelm target resources. The observation that expanding the feature subset beyond eight attributes yields negligible accuracy improvements (and occasionally marginal degradation) empirically validates the Pareto principle in feature engineering: a compact subset of ~17% of the original features captures the vast majority of informative variance required for robust multiclass classification.

The superior performance of the Decision Tree (DT) model (99.97% accuracy with 8 features) underscores the value of interpretable, rule-based classifiers in cybersecurity applications. Unlike black-box deep learning architectures, DT provides explicit decision boundaries that facilitate forensic analysis, compliance auditing, and rapid model updates by security analysts. While Random Forest (RF) achieved comparable results, its ensemble nature introduced slight performance variance on minority classes (e.g., DDoS-Tcp_Flood), likely due to the averaging effect of bagging that dilutes sensitivity to subtle attack signatures. Conversely, the catastrophic degradation of Naïve Bayes (NB) in multiclass settings (accuracy ~30%) confirms the theoretical limitations imposed by its conditional independence assumption. Network flow features extracted from TCP/IP stacks are inherently correlated, particularly under coordinated DDoS campaigns; thus, NB's probabilistic formulation struggles to disentangle overlapping attack manifolds. However, NB's recovery in binary classification (99.82% accuracy) reveals its residual utility for coarse-grained anomaly filtering, where distinguishing benign from malicious traffic requires fewer discriminative dimensions.

The comparative analysis between IG and Principal Component Analysis (PCA) further highlights the importance of preserving semantic integrity in IDS design. While PCA effectively compresses variance into orthogonal components, it obscures the physical meaning of network attributes, which can impair model interpretability and hinder the deployment of PCA-augmented NB models (as evidenced by a high FPR of 0.0310 and a precision of 0.5124 for benign traffic). IG, by contrast, retains original feature semantics, enabling transparent threshold tuning and direct mapping to network protocol behaviors. The hybrid IG-PCA approach demonstrated marginal gains, suggesting that sequential filtering (IG for relevance, PCA for redundancy removal) may benefit high-dimensional

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

datasets with complex feature collinearity. However, the added computational overhead is unjustified for binary DDoS detection, where IG alone suffices.

Contextualizing these findings within the broader literature, recent studies on the CICIoT2023 dataset have predominantly relied on deep learning or wrapper-based selection, often reporting accuracy metrics above 99% but at the cost of significant training latency and memory footprint. Our results indicate that a carefully curated filter-based pipeline can achieve parity in detection performance while reducing computational complexity by 30–40%, making it highly suitable for resource-constrained cloud-IoT edge nodes. This aligns with emerging paradigms advocating for "lightweight AI" in operational technology (OT) and IoT security, where inference speed, energy efficiency, and model transparency are as critical as raw accuracy.

Nevertheless, this study acknowledges several limitations. First, the evaluation relies on a static train-test split without k-fold cross-validation, which may overestimate generalization performance on unseen network topologies or evolving attack patterns. Second, the CICIoT2023 dataset, while comprehensive, reflects lab-generated traffic distributions; real-world cloud-IoT environments exhibit higher concept drift, encrypted traffic, and adaptive adversarial behaviors that were not simulated. Third, the feature selection threshold was determined empirically; adaptive IG cutoff mechanisms or cost-sensitive learning strategies could further optimize the accuracy-efficiency trade-off. Finally, the study focuses exclusively on supervised learning, leaving unsupervised or semi-supervised anomaly detection frameworks unexplored.

Future research should address these gaps by integrating streaming learning architectures that enable continuous model adaptation, evaluating adversarial robustness to evasion attacks (e.g., feature perturbations or traffic mimicry), and deploying the proposed IG-DT pipeline on physical edge devices to measure real-time latency, memory consumption, and power efficiency. Additionally, exploring hybrid frameworks that combine IG-based feature selection with lightweight neural architectures (e.g., 1D-CNNs or temporal convolutional networks) could enhance the detection of encrypted or polymorphic DDoS variants while maintaining computational frugality.

5. CONCLUSION

This study presents a lightweight, interpretable intrusion detection framework for identifying DDoS attack variants in cloud-IoT environments, leveraging Information Gain (IG) for feature selection and machine learning classifiers for both multiclass and binary detection. Using the CICIoT2023 dataset, we demonstrate that an IG-optimized Decision Tree model utilizing only eight highly discriminative features (IAT, Header_Length, Tot_size, Min, Magnitude, AVG, Tot_sum, and Max) achieves a superior multiclass accuracy of 99.97%, while reducing computational complexity by approximately 30–40% compared to full-feature baselines. The systematic evaluation reveals that expanding the feature subset beyond eight attributes yields diminishing returns, reinforcing the principle that feature quality outweighs quantity in efficient IDS design. Furthermore, while Naïve Bayes exhibits severe performance degradation in multiclass scenarios due to violated conditional independence assumptions, it remains highly effective for binary DDoS filtering, underscoring the importance of aligning algorithmic assumptions with task complexity. Comparative analysis with PCA-based extraction confirms that entropy-driven selection preserves the integrity of semantic features, enabling transparent, auditable decision boundaries that are crucial for operational cybersecurity.

The proposed framework offers a practical, resource-efficient solution for real-time DDoS monitoring in cloud-IoT infrastructures, particularly where computational constraints, model interpretability, and rapid deployment are prioritized. Limitations include reliance on static dataset partitions and the absence of adversarial or streaming evaluation. Future work will focus on implementing continuous learning pipelines, testing adversarial robustness, and validating the model on

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

hardware-constrained edge devices to bridge the gap between theoretical performance and operational deployment. Ultimately, this research contributes to the growing paradigm of lightweight, explainable AI for next-generation network security, demonstrating that strategic feature engineering can achieve state-of-the-art detection without compromising efficiency or transparency.

CONFLICT OF INTEREST

The authors declare no conflicts of interest with the research object in this paper.

ACKNOWLEDGEMENT

The authors extend heartfelt gratitude to Universitas Dinamika Bangsa for their outstanding support, which made this research possible. This study was also made possible by the financial assistance from the Direktorat Riset, Teknologi, dan Pengabdian kepada Masyarakat, Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, 2024, whose sponsorship and full support were invaluable.

REFERENCES

[1] P. Borra, "An overview of cloud computing and leading cloud service providers," *Int. J. Comput. Eng. Technol. Vol.*, vol. 15, pp. 122–133, 2024.

[2] M. A. Omer, A. A. Yazdeen, H. S. Malallah, and L. M. Abdulrahman, "A Survey on Cloud Security: Concepts, Types, Limitations, and Challenges," *J. Appl. Sci. Technol. Trends*, vol. 3, no. 02, pp. 47–57, Dec. 2022, doi: 10.38094/jastt301137.

[3] S. Shamkura, "The Impact of Cloud Elasticity and Pay as You Go Pricing on Financial Risk and Cost Optimization for Variable Workloads," *J. Multidiscip.*, vol. 5, no. 7, pp. 56–65, 2025.

[4] A. Mahida, "Comprehensive review on optimizing resource allocation in cloud computing for cost efficiency," *J. Artif. Intell. \& Cloud Comput. SRC/JAICC-249. DOI doi. org/10.47363/JAICC/2022*, vol. 232, pp. 2–4, 2022.

[5] A. Sharma, S. Reddy, P. S. Patwal, D. Gowda, and others, "Data analytics and cloud-based platform for internet of things applications in smart cities," in *2022 International Conference on Industry 4.0 Technology (I4Tech)*, 2022, pp. 1–6.

[6] M. Almutairi and F. T. Sheldon, "IoT--cloud integration security: A survey of challenges, solutions, and directions," *Electronics*, vol. 14, no. 7, p. 1394, 2025.

[7] A. Bhardwaj, V. Mangat, R. Vig, S. Halder, and M. Conti, "Distributed denial of service attacks in cloud: State-of-the-art of scientific and commercial solutions," *Computer Science Review*, vol. 39. Elsevier Ireland Ltd, Feb. 01, 2021. doi: 10.1016/j.cosrev.2020.100332.

[8] P. Verma, N. Bharot, J. G. Breslin, M. Sharma, N. Chaurasia, and A. Vidyarthi, "Uncovering collateral damages and advanced defense strategies in cloud environments against DDoS attacks: A comprehensive review," *Trans. Emerg. Telecommun. Technol.*, vol. 35, no. 4, p. e4934, 2024.

[9] Z. R. Alashhab, M. Anbar, M. M. Singh, I. H. Hasbullah, P. Jain, and T. A. Al-Amiedy, "Distributed Denial of Service Attacks against Cloud Computing Environment: Survey, Issues, Challenges and Coherent Taxonomy," *Applied Sciences (Switzerland)*, vol. 12, no. 23. MDPI, Dec. 01, 2022. doi: 10.3390/app122312441.

[10] L. Poonia and S. Tinker, "A comprehensive analysis of the types, impacts, prevention, and mitigation of DDoS attacks," *Recent Patents Eng.*, vol. 19, no. 9, p. E18722121322166, 2025.

[11] S. Basuli and M. Padhya, "Botnet-Based DDoS Attack: Automatic Detection, Mitigation, and Real-Time Traffic Filtering in Cloud Environments," in *International Conference on Advanced Network Technologies and Intelligent Computing*, 2024, pp. 140–161.

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

[12] A. Odeh, A. Aboshgifa, and N. Belhaj, "Mitigating DDoS attacks in cloud computing environments: Challenges and strategies," in *2023 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, 2023, pp. 1–4.

[13] M. Agoramoorthy, A. Ali, D. Sujatha, M. R. T. F, and G. Ramesh, "An Analysis of Signature-Based Components in Hybrid Intrusion Detection Systems," in *2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS)*, 2023, pp. 1–5. doi: 10.1109/ICCEBS58601.2023.10449209.

[14] S. Balasubramaniam *et al.*, "Optimization enabled deep learning-based DDoS attack detection in cloud computing," *Int. J. Intell. Syst.*, vol. 2023, no. 1, p. 2039217, 2023.

[15] Y. Sanjalawe and T. Althobaiti, "DDoS Attack Detection in Cloud Computing Based on Ensemble Feature Selection and Deep Learning," *Comput. Mater. \& Contin.*, vol. 75, no. 2, 2023.

[16] E. S. GSR, R. Ganeshan, I. D. J. Jingle, and J. P. Ananth, "FACVO-DNFN: Deep learning-based feature fusion and Distributed Denial of Service attack detection in cloud computing," *Knowledge-Based Syst.*, vol. 261, p. 110132, 2023.

[17] A. Berguiga, A. Harchay, and A. Massaoudi, "HIDS-IoMT: A Deep Learning-Based Intelligent Intrusion Detection System for the Internet of Medical Things," *IEEE Access*, vol. 13, pp. 32863–32882, 2025, doi: 10.1109/ACCESS.2025.3543127.

[18] A. H. Halbouni, T. S. Gunawan, M. Halbouni, F. A. A. Assaig, M. R. Effendi, and N. Ismail, "CNN-IDS: Convolutional neural network for network intrusion detection system," in *2022 8th International Conference on Wireless and Telematics (ICWT)*, 2022, pp. 1–4.

[19] E. U. H. Qazi, M. H. Faheem, and T. Zia, "HDLNIDS: a hybrid deep-learning-based network intrusion detection system," *Appl. Sci.*, vol. 13, no. 8, p. 4921, 2023.

[20] V. Sharma, A. Rai, Y. Dixit, Y. Tomar, R. Rai, and T. Sharma, "Interpretable Deep Learning Framework for Anomaly Detection in High-Dimensional Network Traffic Data," *Reconstruction*, vol. 26, no. 30, p. 40.

[21] C. Kavitha, S. M, T. R. Gadekallu, N. K, B. P. Kavin, and W.-C. Lai, "Filter-based ensemble feature selection and deep learning model for intrusion detection in cloud computing," *Electronics*, vol. 12, no. 3, p. 556, 2023.

[22] G. N. Tikhe and P. S. Patheja, "A wrapper feature selection based hybrid deep learning model for DDoS detection in a network with NFV behaviors," *Wirel. Pers. Commun.*, vol. 133, no. 1, pp. 481–506, 2023.

[23] S. V. Dicholkar and J. H. Nirmal, "DoS Attack Detection Using Feature Selection with Information Gain and ML Classification," in *2024 Second International Conference on Advances in Information Technology (ICAIT)*, 2024, pp. 1–6.

[24] M. Alduailij, Q. W. Khan, M. Tahir, M. Sardaraz, M. Alduailij, and F. Malik, "Machine-Learning-Based DDoS Attack Detection Using Mutual Information and Random Forest Feature Importance Method," *Symmetry (Basel)*, vol. 14, no. 6, pp. 1–15, 2022, doi: 10.3390/sym14061095.

[25] A. A. Elshweikh, A. M. Maher, M. Hussein, and A. D. Elbayoumy, "Intrusion Detection System for IoT Using CICIoT2023 Dataset," in *2024 6th Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, 2024, pp. 512–516.

[26] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, and A. A. Ghorbani, "CICIoT2023: A Real-Time Dataset and Benchmark for Large-Scale Attacks in IoT Environment," *Sensors*, vol. 23, no. 13, 2023, doi: 10.3390/s23135941.

[27] W. A. H. Salman and C. H. Yong, "Overview of the CICIoT2023 Dataset for Internet of Things Intrusion Detection Systems," *Mesopotamian J. Big Data*, vol. 2025, pp. 50–60, 2025, doi:

Vol. 7, No. 3, June 2026, Page. 2994-3011
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5752

10.58496/MJBD/2025/004.

[28] A. M. Al-Ghamdi and M. M. Alansari, "Enhancing IoT Security: A Comparative Study of CNN and RNN-Based Anomaly Detection Using the CICIoT2023 Dataset," *IAENG Int. J. Comput. Sci.*, vol. 52, no. 5, 2025.

[29] E. Halabaku and E. Bytyçi, "Overfitting in Machine Learning: A Comparative Analysis of Decision Trees and Random Forests," *Intell. Autom. \& Soft Comput.*, vol. 39, no. 6, 2024.

[30] J. Zhang, "A Random Forest-Based Approach for Cybersecurity Attack Detection," *J. Next Comput.*, vol. 1, no. 1, pp. 1–12, 2025.

[31] K. H. Le, M. H. Nguyen, T. D. Tran, and N. D. Tran, "IMIDS: An Intelligent Intrusion Detection System against Cyber Threats in IoT," *Electron.*, vol. 11, no. 4, pp. 1–16, 2022, doi: 10.3390/electronics11040524.

[32] S. H. Mohammed *et al.*, "A review on the evaluation of feature selection using machine learning for cyber-attack detection in smart grid," *IEEE Access*, vol. 12, pp. 44023–44042, 2024.

[33] N. Hasdyna and R. K. Dinata, "A hybrid optimization of supervised learning models using information gain-based feature selection," *Int. J. Comput.*, vol. 24, no. 1, pp. 178–189, 2025.

[34] M. R. Islam, A. A. Lima, S. C. Das, M. F. Mridha, A. R. Prodeep, and Y. Watanobe, "A comprehensive survey on the process, methods, evaluation, and challenges of feature selection," *IEEE Access*, vol. 10, pp. 99595–99632, 2022.

[35] N. Uddamari and P. Sammulal, "Ensemble-Based Network Anomaly Detection Using RFE and Information Gain for Optimized Feature Selection," *Informatica*, vol. 49, no. 10, 2025.

[36] A. S. Afolabi and O. A. Akinola, "Network intrusion detection using knapsack optimization, mutual information gain, and machine learning," *J. Electr. Comput. Eng.*, vol. 2024, no. 1, p. 7302909, 2024.

[37] R. Tekin, O. Yaman, and T. Tuncer, "Decision Tree Based Intrusion Detection Method in the Internet of Things," *Int. J. Innov. Eng. Appl.*, vol. 6, no. 1, pp. 17–23, 2022.

[38] Z. Sun, G. Wang, P. Li, H. Wang, M. Zhang, and X. Liang, "An improved random forest based on the classification accuracy and correlation measurement of decision trees," *Expert Syst. Appl.*, vol. 237, p. 121549, 2024.

[39] H. A. Salman, A. Kalakech, and A. Steiti, "Random Forest Algorithm Overview," *Babylonian J. Mach. Learn.*, vol. 2024, pp. 69–79, 2024, doi: 10.58496/BJML/2024/007.

[40] M. Alduailij, Q. W. Khan, M. Tahir, M. Sardaraz, M. Alduailij, and F. Malik, "Machine-learning-based DDoS attack detection using mutual information and random forest feature importance method," *Symmetry (Basel)*, vol. 14, no. 6, p. 1095, 2022.

[41] A. Abbas, M. A. Khan, S. Latif, M. Ajaz, A. A. Shah, and J. Ahmad, "A New Ensemble-Based Intrusion Detection System for Internet of Things," *Arab. J. Sci. Eng.*, vol. 47, no. 2, pp. 1805–1819, Feb. 2022, doi: 10.1007/s13369-021-06086-5.

[42] N. Konyrbaev *et al.*, "Evaluation and Optimization of The Naive Bayes Algorithm For Intrusion Detection Systems Using The USB-IDS-1 Dataset," *Eastern-European J. Enterp. Technol.*, vol. 132, no. 2, 2024.

[43] S. Naiem, A. E. Khedr, A. M. Idrees, and M. I. Marie, "Enhancing the efficiency of Gaussian Naive Bayes machine learning classifier in the detection of DDOS in cloud computing," *IEEE Access*, vol. 11, pp. 124597–124608, 2023.