Jurnal Informatika: Jurnal pengembangan IT

Vol. 11, No. 2, 2026
ISSN : 2477-5126, DOI:10.30591/jpit.v11i2.10378

Anomaly-Based Network Intrusion Detection Using Isolation Forest on the Imbalanced UNSW-NB15 Dataset

Syifaurachman[1], Samso Supriyatna[2], Muhamad Ihsan Ashari[3]
[1,2,3]Department of Information Systems, Faculty of Computer Science, Universitas Pamulang, Indonesia
[1]dosen03181@unpam.ac.id, [2]dosen02830@unpam.ac.id, [3]dosen03154@unpam.ac.id

Info Artikel

Riwayat Artikel:

Received 2026-03-05
Revised 2026-05-11
Accepted 2026-05-14

Corresponding Author:

Syifaurachman

Email: dosen03181@unpam.ac.id

Abstract *– Modern cyber threats require adaptive intrusion detection systems (IDS) capable of identifying complex anomalies within large networks. . Traditional signature-based methods frequently struggle to detect novel zero-day attacks, making unsupervised machine learning, specifically Isolation Forest (IF), a vital research direction. IF detects anomalies by isolating unusual observations through random partitioning mechanisms.. This study evaluates IF for intrusion detection using the highly imbalanced UNSW-NB15 dataset, comprising 2,540,044 records (87% normal traffic, 13% attack traffic). A quantitative approach was applied, evaluating performance across Accuracy, Precision, Recall, F1-score, ROC-AUC, and PR-AUC. Results show an accuracy of 79.3% and a ROC-AUC of 0.6302. However, a low PR-AUC (0.1710) and F1-score (0.18) reveal the model's critical failure in detecting the minority attack class due to severe feature overlap. These findings prove that relying strictly on Accuracy or ROC-AUC is misleading in imbalanced scenarios. Therefore, IF should serve only as a lightweight baseline screening layer; hybrid architectures are necessary to improve detection sensitivity in real-world IDS.*

Keywords: *Imbalanced Dataset, Intrusion Detection, Isolation Forest, Unsupervised Learning*

Abstrak *– Ancaman siber modern menuntut penerapan sistem deteksi intrusi (IDS) adaptif yang mampu mengidentifikasi anomali kompleks di jaringan berskala besar. Keterbatasan metode signature-based konvensional dalam mendeteksi zero-day attack baru menjadikan machine learning berbasis unsupervised, khususnya algoritma Isolation Forest (IF), sebagai arah riset vital. IF mendeteksi potensi anomali dengan mengisolasi observasi tidak biasa melalui mekanisme partisi acak. Penelitian ini mengevaluasi IF untuk deteksi intrusi menggunakan dataset UNSW-NB15 yang sangat tidak seimbang, mencakup total 2.540.044 rekaman (87% lalu lintas normal, 13% serangan). Pendekatan kuantitatif diterapkan untuk mengevaluasi performa prediksi berdasarkan Akurasi, Precision, Recall, F1-score, ROC-AUC, dan PR-AUC. Hasil eksperimen menunjukkan akurasi 79,3% dan ROC-AUC 0,6302. Meskipun demikian, nilai PR-AUC (0,1710) dan F1-score (0,18) yang rendah jelas mengungkap kegagalan kritis model dalam mendeteksi kelas minoritas akibat tumpang tindih fitur yang parah. Temuan ini membuktikan bahwa bergantung mutlak pada Akurasi atau ROC-AUC sangat menyesatkan pada kondisi imbalanced. Oleh karena itu, IF murni seharusnya hanya berfungsi sebagai lapisan penyaring awal (baseline screening); arsitektur hybrid mutlak diperlukan untuk meningkatkan sensitivitas deteksi di lingkungan IDS dunia nyata.*

Kata Kunci*: Imbalanced Dataset, Intrusion Detection, Isolation Forest, Unsupervised Learning*

I. INTRODUCTION

The advancement of digital infrastructure has increased the complexity of network traffic while simultaneously expanding the cyber attack surface. Signature-based Intrusion Detection Systems (IDS) are effective in identifying known attack patterns; however, they lack adaptability when confronting emerging threats or zero-day attacks. Anomaly detection approaches using unsupervised learning have emerged as a promising alternative for identifying emerging or zero-day threats that lack predefined signatures, as they model normal system behavior and identify deviations as potential indicators of malicious activity [1], [2]. While supervised models like Random Forest demonstrate high performance, this is typically achieved under balanced conditions or through artificial resampling [8]. In real-world operational scenarios characterized by extreme data imbalance, supervised approaches are prone to majority class bias, making the evaluation of unsupervised algorithms on naturally imbalanced datasets crucial.

Isolation Forest, introduced by Liu et al. [3], [4], is an ensemble-based anomaly detection algorithm that isolates observations through random partitioning. Instances that are more easily isolated are more likely to be classified as anomalies. This principle makes Isolation Forest theoretically suitable for handling imbalanced datasets like UNSW-NB15, as anomalies are 'few and different' and can be isolated without requiring artificial data balancing techniques. Furthermore, the algorithm exhibits linear time complexity $O(n)$, providing the necessary scalability to process the 2,540,044 records efficiently [5].

The UNSW-NB15 dataset [6] has become a widely adopted benchmark, as it reflects contemporary network traffic and includes nine attack categories. However, the class distribution within this dataset is imbalanced, with a predominance of normal traffic instances. Such extreme imbalance mathematically leads to biased performance evaluations if accuracy is used as the sole metric, a phenomenon well-documented in imbalanced learning literature where minority class failures are often masked by the sheer volume of the majority class. Furthermore, network intrusion detection utilizing the UNSW-NB15 dataset has gained increasing empirical attention precisely because it accurately reflects the growing complexity and diversity of modern cyber attacks, rendering older benchmark datasets obsolete. Consequently, developing more effective detection mechanisms that can navigate this data imbalance while maintaining low false positive rates has become a critical requirement in contemporary network security environments. Additionally, the rapid rise of cyber threats highlights the need for advanced intrusion detection techniques capable of achieving high detection accuracy while maintaining low false positive rates [7].

Supervised models such as Random Forest and deep learning techniques have traditionally demonstrated high performance when labeled data are available, particularly under balanced class distributions or when artificial resampling techniques are applied [8]. However, several studies report critical limitations of such machine learning–based intrusion detection systems when applied to the UNSW-NB15 dataset under naturally imbalanced conditions. These limitations are specifically related to class overlap and the inherent overfitting and skewness sensitivity of Random Forest, which heavily biases the model toward the majority class and severely degrades real-world operational performance [9].

Furthermore, despite reporting high accuracy and sensitivity, previous work has not sufficiently examined the trade-off between sensitivity and specificity, cross-dataset generalizability, or misclassification among attack classes. The mutual information–based feature selection method improves intrusion detection by selecting optimal features, leading to higher accuracy and lower computational cost [10].

Intrusion Detection Systems (IDS) face challenges in detecting new attack types because models are typically trained on specific datasets, while evolving attack strategies require more adaptive detection approaches, to address this limitation and the risk of underfitting or overfitting in conventional machine learning methods, the study proposes an accurate IDS model using multiple datasets and a hybrid feature selection technique [11]. Outlier detection in time series data remains a critical challenge because existing approaches often experience overfitting, which reduces detection effectiveness. To overcome this limitation, the study introduces ensemble frameworks using sparsely connected recurrent autoencoder networks to improve detection performance [12]. The study highlights the importance of improving network intrusion detection using the UNSW-NB15 dataset, as the traditional KDD'99 dataset has become outdated for modern network security analysis. To address this gap, the research explores ensemble learning methods, including Random Forest and XGBoost, to improve prediction accuracy in intrusion detection tasks [13]. The study proposes an Isolation Forest–based anomaly detection model for Industrial Control Systems using the HAI dataset, achieving improved detection accuracy, lower false positives, and better computational efficiency than traditional methods [14]. The SPIF algorithm improves the scalability and efficiency of network traffic anomaly detection by leveraging Spark-based parallel processing, demonstrating faster computation and effective performance on the UNSW-NB15 dataset compared to traditional Isolation Forest approaches [15].

The study [16] proposes a hybrid intrusion detection framework combining denoising autoencoder feature learning, Isolation Forest anomaly scoring, and LightGBM classification with SMOTE-ENN balancing, achieving high detection accuracy and strong performance for rare attack classes on the NSL-KDD and UNSW-NB15 datasets. The study introduces a deep learning–based framework for detecting rare anomalies in highly imbalanced network traffic by integrating Isolation Forest and SMOTE-Tomek for data balancing, along with a Residual Convolutional Autoencoder and Bidirectional GRU, resulting in high anomaly detection accuracy and performance [17]. The study [18] investigates anomaly detection in industrial environments using one-class classification techniques that learn from legitimate network traffic, while addressing the lack of interpretability in deep learning–based OCC models for effective and lightweight intrusion detection systems. A hybrid anomaly detection framework, HDBIF-CO, is introduced by integrating Density-Based Isolation Forest with Coati Optimization, along with preprocessing, feature selection, and DBSCAN clustering to enhance cyberattack detection on the NSL-KDD, CICIDS2017, and UNSW-NB15 datasets [19]. Cybersecurity and privacy risks in IoT and IIoT environments within Industry 4.0 require more effective intrusion detection approaches due to the complexity and data-intensive nature of interconnected systems. To address this challenge, an anomaly-based intrusion detection system using a voting-based ensemble model (ABIDS-VEM) is proposed to improve detection accuracy and efficiency [20]. An optimized intrusion detection framework integrating data sampling, feature selection using Random Forest and Genetic Algorithm, and anomaly detection with Isolation Forest is proposed to improve detection accuracy on the UNSW-NB15 dataset. Experimental results show that the optimized XGBGA model achieves superior performance with an F1-score of 92.8% and AUC of 97%, effectively detecting rare cyber threats and reducing false alarm rates [21]. An adaptive intrusion detection framework is proposed to enhance cloud security in fog and edge computing environments by addressing IoT data management challenges and improving the detection of various cyber attacks [22]. A honeypot and blockchain-based intrusion detection and prevention framework (HB-IDP) is introduced for IoT environments, integrating improved

Isolation Forest and ensemble learning to detect and prevent cyberattacks using the UNSW-NB15 and BoT-IoT datasets while reducing latency through edge computing [23]. Security vulnerabilities in the Industrial Internet of Things (IIoT) require more effective intrusion detection approaches, as traditional IDS methods often lack integrated feature engineering and machine learning techniques. To address this issue, an IDS model combining Isolation Forest and Pearson's Correlation Coefficient is proposed to improve detection accuracy while reducing computational cost [24]. Intrusion detection in Industry 4.0 environments requires more robust anomaly detection methods to ensure the security and reliability of industrial systems. To address this challenge, the UInDeSI4.0 approach is proposed to enhance anomaly detection using advanced learning techniques on industrial datasets [25]. A fusion-based anomaly detection approach using a modified Isolation Forest is proposed for IoT network intrusion detection, demonstrating improved accuracy, detection rate, and reduced runtime when evaluated on benchmark datasets such as UNSW-NB15 and NSL-KDD[26]. Anomaly detection in multiaspect data streams remains challenging due to correlation complexity and the limitations of dimensionality reduction–based methods. To address this issue, the MDSAD approach is proposed to improve anomaly detection performance, scalability, and group anomaly detection capability [27]. The challenge of extreme data imbalance in modern cyberattack detection has also been a major focus in recent literature published within regional contexts. For instance, recent studies have demonstrated that addressing class imbalance is critically important to optimize the detection accuracy of machine learning algorithms against complex network anomalies, such as attacks in Internet of Vehicles (IoV) environments [28].

Nevertheless, in operational environments, labeling network traffic data requires substantial resources and is not always feasible. Consequently, unsupervised approaches have become increasingly relevant in the context of modern intrusion detection. Based on these considerations, this study focuses on evaluating the performance of Isolation Forest under imbalanced class distribution conditions using evaluation metrics that are more representative of minority classes. This research contributes by providing a realistic quantitative assessment of the effectiveness of an unsupervised model within modern IDS scenarios. Despite extensive research assessing machine learning methods for intrusion detection, a significant research gap still remains the comprehensive behavioral examination of entirely unsupervised models on large-scale datasets such as UNSW-NB15, absent the influence of artificial data balance. This research contributes a critical quantitative analysis of evaluation metric bias, specifically demonstrating the misleading characteristics of ROC-AUC in contrast to PR-AUC regarding minority class failures, and elucidating the inherent operational risk (False Negatives) associated with utilizing a static-threshold Isolation Forest in a naturally imbalanced network environment.

II. METHOD

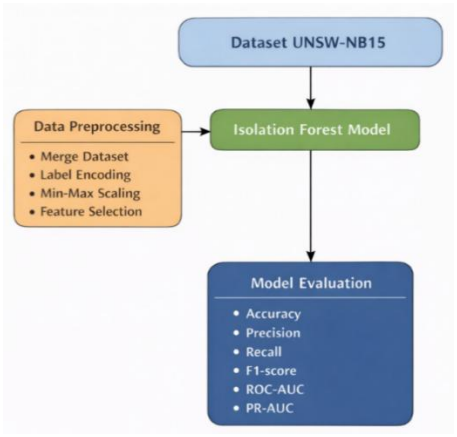

Figure 1. Researh Method

This study adopts a quantitative experimental approach to evaluate the performance of the Isolation Forest algorithm on the UNSW-NB15 dataset under imbalanced class distribution conditions. All experimental stages were explicitly designed to evaluate the model under naturally imbalanced class distribution conditions typical of real-world network traffic, deliberately avoiding artificial resampling techniques. This offline evaluation focuses on structural data imbalance rather than real-time deployment factors such as data streaming or concept drift. The overall research workflow is illustrated in Figure 1, which outlines the sequential stages of dataset acquisition, data preprocessing, model development, and performance evaluation.

The dataset utilized in this research is UNSW-NB15, developed by Moustafa and Slay [6]. The dataset comprises 2,540,044 records with 49 features representing characteristics of modern network traffic, including nine attack categories and normal traffic. The class distribution exhibits a significant dominance of normal traffic

(approximately 87%) compared to attack instances (approximately 13%). This distribution was preserved without applying resampling techniques to realistically represent operational network conditions.

The data preprocessing stage was conducted to ensure data consistency and readiness prior to model training. This process included merging four dataset subsets into a unified dataset, transforming categorical features using Label Encoding, and normalizing numerical features through the Min-Max Scaling method. Additionally, feature selection was performed based on relevance and correlation analysis to reduce redundancy and improve model stability. No oversampling or undersampling techniques were applied in order to maintain the original class distribution characteristics.

In this study, the Isolation Forest algorithm was developed using the Python programming language and the Scikit-learn machine learning framework. The parameters employed in this study included 300 trees (n_estimators = 300) and the default value for max_samples (auto). To establish a clear decision boundary for evaluation, the contamination parameter was set to approximately 0.13. It is crucial to clarify that this value was utilized strictly as a global threshold to binarize the final anomaly scores, reflecting an assumed operational expectation of malicious traffic volume (based on historical network profiling), rather than utilizing ground-truth labels to inform the model. The Isolation Forest algorithm itself constructed its partitions entirely without access to individual class labels, thereby strictly preserving the integrity of the unsupervised learning process as described in [3], [4].

Unlike supervised machine learning methodologies, this study adopts a whole-dataset evaluation strategy, which is a standard approach for strictly unsupervised anomaly detection tasks. The entire dataset of over 2.5 million records was utilized to construct the Isolation Forest model (fit), and the same dataset was subsequently used to generate anomaly predictions (predict). Because the algorithm does not rely on class labels to learn decision boundaries, this approach evaluates the inherent capability of the model to map the structural distribution of the data and isolate deviations without suffering from traditional label-based overfitting.

Furthermore, the threshold decision process was governed directly by the contamination parameter, which was set to approximately 0.1265, mirroring the actual proportion of attack instances in the dataset. Internally, the Isolation Forest computes an anomaly score for each data point based on its average path length across the forest. The decision threshold is mathematically defined as the cut-off point that isolates the top 12.65% of instances with the most extreme anomaly scores, subsequently classifying them as attacks (anomalies).

Model performance was evaluated using multiple quantitative metrics, including Accuracy, Precision, Recall, F1-score, Receiver Operating Characteristic – Area Under the Curve (ROC-AUC), and Precision-Recall Area Under the Curve (PR-AUC). The PR-AUC metric was specifically employed due to its higher sensitivity to class imbalance compared to ROC-AUC. The use of this combination of evaluation metrics aims to provide a more comprehensive assessment of the model's ability to detect anomalies under imbalanced data conditions.

III. RESULT AND DISCUSSION

A. Class Distribution Analysis and Its Implications

The class distribution in the UNSW-NB15 dataset reveals a significant imbalance, comprising 2,540,044 total records, of which approximately 87% correspond to normal traffic and 13% to attack traffic, as illustrated in Figure 2. This distribution reflects realistic network operational conditions, where benign activities substantially outnumber malicious ones. However, from a machine learning perspective, such imbalance may introduce bias toward the majority class.

In absolute terms, the dataset contains more than two million normal instances, while attack instances represent only a small fraction of the total observations. As shown in Figure 2, this disproportionate distribution creates a scenario in which a model may achieve seemingly high accuracy by predominantly predicting instances as normal traffic. Consequently, relying solely on accuracy as an evaluation metric would provide a misleading assessment of the model's ability to detect minority attack events.

This phenomenon has been widely discussed in the anomaly detection and imbalanced data literature [1], [9], which emphasizes the importance of employing complementary metrics to mitigate biased performance interpretation. By preserving the original class distribution without applying resampling or balancing techniques, this study intentionally simulates realistic deployment conditions. As depicted in Figure 2, the dominance of normal traffic presents a challenging detection environment, thereby providing a more operationally meaningful evaluation of Isolation Forest as an unsupervised anomaly detection method.

```
# sklearn: -1 = anomaly, 1 = normal
y_pred_raw = model.predict(X)

# Ubah ke format kita: 1 = attack/anomaly, 0 = normal
y_pred = np.where(y_pred_raw == -1, 1, 0)

print("Total data:", len(y_pred))
print("Total predicted anomaly:", sum(y_pred))
print("Persentase anomaly:", sum(y_pred)/len(y_pred))
```

```
Total data: 2540047
Total predicted anomaly: 321283
Persentase anomaly: 0.12648702957071267
```

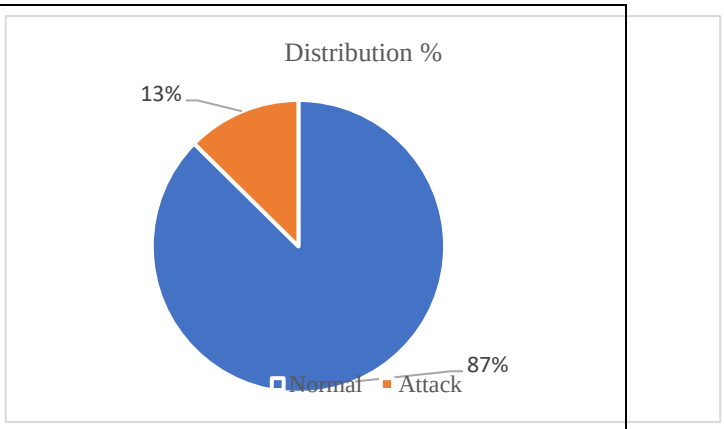

Figure 1. Data Distribution Normal vs Attack

To provide a clearer context regarding the dataset's structural characteristics, Table 2 presents a representative sample of the raw network traffic features utilized in this study.

TABLE 1.
EXAMPLE OF NORMAL AND ATTACK TRAFFIC FEATURES IN UNSW-NB15

id	srcip	sport	dstip	dsport	proto	state	…	attack_cat	Label
0	59.166.0.9	7045	149.171.126.7	25	tcp	FIN	…	NaN	0
1	59.166.0.9	9685	149.171.126.2	80	tcp	FIN	…	NaN	0
2	59.166.0.2	1421	149.171.126.4	53	udp	CON	…	NaN	0
3	59.166.0.2	21553	149.171.126.2	25	tcp	FIN	…	NaN	0
…							…	NaN	0
440043	175.45.176.0	17293	149.171.126.17	110	ccp	CON	…	Exploits	1

As illustrated in Table 1, while there are variations in packet transmission behaviors, sophisticated attacks often mimic the characteristics of normal traffic, complicating the detection process for unsupervised models.

B. Confusion Matrix Analysis and Its Impact on IDS

The confusion matrix results indicate the presence of both false positives and false negatives in non-negligible quantities, as illustrated in Figure 3, while the detailed classification metrics are summarized in Table 2. From the confusion matrix, the model correctly classifies the majority of normal traffic instances (True Negatives = 1,956,142). However, it also produces approximately 262,672 False Positives (FP) and 206,622 False Negatives (FN), with only 56,061 True Positives (TP) detected for the attack class. Using these values, the False Positive Rate (FPR) and False Negative Rate (FNR) can be explicitly calculated:

$$FPR = \frac{FP}{FP+TN} \quad (1)$$

$$FNR = \frac{FN}{FN+TP} \quad (2)$$

Substituting the observed values:

$$FPR \approx \frac{262{,}672}{262{,}672 + 1{,}956{,}142} \approx 0.1184 \approx 11.84\% \quad (3)$$

This indicates that approximately 11.84% of normal traffic is incorrectly flagged as malicious, which may contribute to alert fatigue in operational environments.

TABLE 2.
CLASSIFICATION REPORT

data	*precision*	*recall*	*f1-score*	*support*
0	0.88	0.88	0.88	2218764
1	0.18	0.18	0.18	321283
accuracy			0.79	2540047
macro avg	0.53	0.53	0.53	2540047
weighted avg	0.79	0.79	0.79	2540047

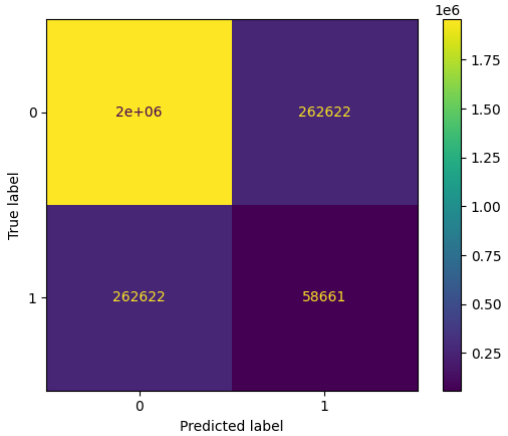

Figure 2. Confusion Matrix Results

Meanwhile,

$$FNR \approx \frac{206{,}622}{206{,}622 + 56{,}061} \approx 0.8173 \approx 81.73\% \quad (4)$$

This implies that exactly 81.73% of attack instances are not detected, highlighting a critical failure in the model's sensitivity toward the minority class. In the context of an Intrusion Detection System (IDS), this magnitude of false negatives carries severe operational consequences, as over 206,000 malicious activities would bypass the security perimeter undetected. Furthermore, the claim of 'acceptable performance' on the majority class cannot be justified; generating 262,672 False Positives (an 11.84% error rate on normal traffic) would inevitably cause massive alert fatigue for security analysts in a real-world implementation, rendering the system highly inefficient.

Given the massive volume of misclassifications, we hypothesize that the feature distributions of normal and sophisticated mimicry attacks in the UNSW-NB15 dataset heavily overlap. Because Isolation Forest relies purely on isolation depth to detect anomalies, this presumed overlap severely reduces the geometric separability between classes. While this study does not provide empirical feature space visualizations (such as PCA or t-SNE) to map this overlap, the magnitude of the model's failure strongly suggests that statistical isolation alone cannot distinguish deeply embedded malicious patterns. Future studies must empirically verify this hypothesis through in-depth feature space analysis.

C. ROC-AUC Evaluation: Discriminative Capability of the Model

The obtained ROC-AUC value falls within a moderate range. ROC-AUC measures the model's ability to discriminate between two classes across various threshold settings. A moderate score indicates that Isolation Forest possesses reasonable discriminative capability but is not fully optimal in separating all attack instances from normal traffic. Theoretically, Isolation Forest relies on isolation depth as an anomaly indicator. In datasets with complex distributions or overlapping feature spaces, the difference in isolation path lengths between normal and attack classes becomes less distinct. This phenomenon may lead to a decrease in AUC values.

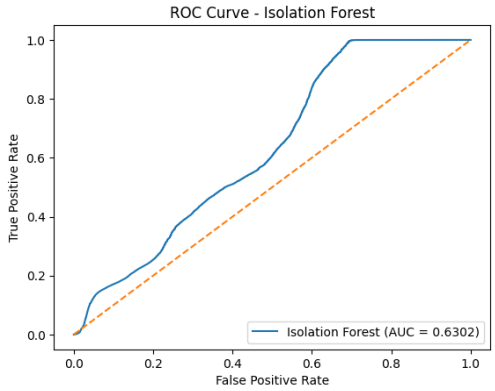

Figure 3. ROC Curve Result Isolation Forest with UNSW-NB15 imbalance dataset

The Isolation Forest model achieved an ROC-AUC value of 0.6302 (as illustrated in Figure 4). While statistically this value sits above the random guessing baseline of 0.5, characterizing it as "moderate" in the context of an Intrusion Detection System (IDS) is highly misleading. Given the critically low recall of 18% and the massive False Negative Rate of 81.73% observed in the confusion matrix, an ROC-AUC of 0.6302 actually reflects a weak operational performance.

This apparent paradox—achieving an ROC-AUC above 0.6 while failing to detect the vast majority of attacks—lies at the core of the evaluation bias inherent in highly imbalanced datasets. The Receiver Operating Characteristic (ROC) curve evaluates the trade-off between the True Positive Rate (TPR) and the False Positive Rate (FPR). The FPR calculation incorporates True Negatives (TN) in its denominator. Because the UNSW-NB15 dataset is dominated by over 2.2 million normal instances, the model's ability to correctly classify a large portion of them artificially inflates the ROC-AUC score. The sheer volume of True Negatives exerts a mathematically disproportionate influence on the curve, effectively masking the model's critical failure in detecting the minority attack class.

Regarding the performance drivers behind this failure, we hypothesize that the feature distributions of normal traffic and sophisticated attacks heavily overlap within the dataset. Isolation Forest identifies anomalies purely based on structural isolation depth through random partitioning. When attack vectors are specifically designed to mimic normal traffic patterns, they embed themselves deeply within the normal data distribution, rendering the isolation mechanism ineffective. While this study does not provide empirical feature space visualizations (such as PCA or t-SNE) to visually map this overlap, the magnitude of the model's false negatives strongly supports this hypothesis. Future research must empirically verify this overlap through dedicated feature space analysis.

Furthermore, while supervised classification models trained on labeled data can explicitly optimize decision boundaries to separate such overlapping classes [8], unsupervised models like Isolation Forest lack this advantage. However, the most critical takeaway from this analysis is not merely the limitation of the algorithm, but the deceptive nature of the ROC-AUC metric itself. A seemingly "adequate" ROC-AUC creates a dangerous false sense of security in cybersecurity deployments. This structural evaluation bias definitively proves that in imbalanced intrusion detection environments, the evaluation must shift away from ROC-AUC and prioritize Precision-Recall Area Under the Curve (PR-AUC) as the primary metric to expose a model's true capability against minority threats.

D. Precision-Recall Evaluation and Sensitivity to Minority Class

The Precision-Recall (PR) curve shown in Figure 5 provides a more informative evaluation under imbalanced data conditions. The obtained PR-AUC value of 0.1710 indicates limited effectiveness of the model in accurately identifying minority-class (attack) instances. Unlike ROC-AUC, which evaluates the trade-off between True Positive Rate and False Positive Rate, PR-AUC directly measures the balance between precision and recall for the positive class.

In imbalanced datasets such as UNSW-NB15, PR-AUC is particularly important because it reflects how well the model performs on the minority attack class without being influenced by the large volume of normal traffic. A PR-AUC value close to the attack prevalence rate suggests that the model provides only marginal improvement over random classification. Given that attack instances represent approximately 13% of the dataset, the PR-AUC value of 0.1710 indicates that the model performs only slightly above a random baseline.

This result aligns with the confusion matrix findings presented in Section 3.2, where the recall for the attack class was approximately 18% and the False Negative Rate reached 81.7%. The PR curve in Figure 5 demonstrates that precision declines rapidly as recall increases, indicating difficulty in maintaining a stable balance between detecting more attacks and avoiding false alarms.

Compared to the ROC-AUC value of 0.6302 (Section 3.3), the significantly lower PR-AUC confirms that ROC-based evaluation may present an overly optimistic assessment in imbalanced scenarios. As highlighted in prior studies [9], PR-AUC offers a more realistic performance interpretation when the positive class is rare.

Therefore, the PR-AUC analysis reinforces the conclusion that while Isolation Forest retains moderate ranking capability, its threshold-based detection effectiveness for minority attack instances remains limited. This finding underscores the importance of selecting evaluation metrics that appropriately reflect operational IDS requirements under imbalanced conditions.

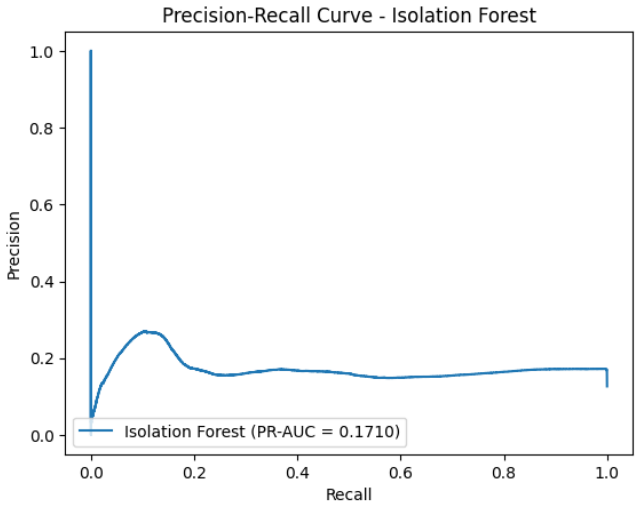

Figure 4. PR Curve Result Isolation Forest with UNSW-NB15 imbalance dataset

E. Practical Implications for IDS Implementation

The results of this study indicate that Isolation Forest can function as a baseline anomaly detection mechanism in modern network environments. However, its deployment must be carefully contextualized within the limitations observed under imbalanced data conditions. While the model achieved an overall accuracy of approximately 79.3% and a moderate ROC-AUC of 0.6302, the low recall for the attack class (18%) and high False Negative Rate (81.7%) highlight a critical gap in minority-class detection capability.

In real-world implementations, an unsupervised IDS such as Isolation Forest is more appropriately positioned as an initial anomaly screening layer rather than a standalone detection solution. Its strength lies in ranking and filtering potentially anomalous traffic from large volumes of normal data without requiring labeled instances. However, given the observed detection limitations, especially for minority attack classes, it should be integrated with secondary analysis mechanisms, such as supervised classifiers, signature-based detection modules, or human-in-the-loop validation processes.

Adaptive thresholding strategies may improve detection sensitivity by dynamically adjusting the anomaly score cutoff based on operational risk tolerance. Additionally, hybrid architectures—combining unsupervised anomaly detection with supervised classification—may reduce false negatives while maintaining manageable false positive rates. These findings also emphasize that evaluation metrics significantly influence the interpretation of IDS effectiveness. Although accuracy and ROC-AUC may suggest moderate performance, PR-AUC (0.1710) reveals the model's limited capability in detecting rare attack events. Therefore, reliance on accuracy-based evaluation alone may lead to overly optimistic and potentially misleading conclusions in imbalanced intrusion detection scenarios. From a practical standpoint, this study underscores the necessity of aligning evaluation metrics, deployment strategy, and risk management objectives when implementing machine learning-based IDS in operational network environments.

F. Discussions

The findings of this study highlight the inherent challenges of applying unsupervised anomaly detection methods to imbalanced intrusion detection datasets such as UNSW-NB15. Although the Isolation Forest model achieved an overall accuracy of approximately 79.3% and a moderate ROC-AUC of 0.6302, deeper metric analysis reveals significant limitations in detecting minority attack instances. The low recall (18%) and high False Negative Rate (81.7%) indicate that the model fails to identify a substantial portion of malicious traffic under realistic class distribution conditions.

This discrepancy between accuracy and minority-class performance reflects the well-known accuracy paradox in imbalanced learning scenarios. While the model demonstrates strong capability in recognizing dominant normal traffic, its discriminative boundary remains insufficiently robust for rare attack patterns. The PR-AUC value of 0.1710 further confirms that performance evaluation based solely on ROC metrics may lead to overly optimistic interpretations in skewed datasets.

From a methodological perspective, the results emphasize the structural limitation of Isolation Forest when feature overlap exists between normal and attack traffic. Since the algorithm relies on isolation depth rather than

explicit class boundary optimization, overlapping distributions reduce separability and consequently degrade detection sensitivity. This limitation becomes more pronounced in datasets with high class imbalance.

Despite these constraints, the study confirms that Isolation Forest retains practical value as a lightweight anomaly ranking mechanism. Its independence from labeled data makes it suitable for environments where attack labeling is costly or incomplete. However, for operational IDS deployment, the model should be complemented by adaptive thresholding, hybrid supervised integration, or multi-stage detection architectures to mitigate high false negative rates.

Overall, this study contributes to a more realistic understanding of unsupervised anomaly detection performance under imbalanced conditions. Rather than reporting inflated accuracy metrics, the comprehensive evaluation framework adopted in this research provides a balanced and operationally relevant assessment of model effectiveness in modern network security contexts.

Although this study provides a realistic evaluation of Isolation Forest under naturally imbalanced conditions, several limitations should be acknowledged. The original class distribution of the UNSW-NB15 dataset was preserved without applying resampling strategies such as oversampling or undersampling. While this design reflects operational network environments, it may limit sensitivity toward minority attack instances. In addition, the study focuses exclusively on a single unsupervised algorithm without comparative benchmarking against supervised or hybrid models. Hyperparameter tuning was not systematically optimized, and the evaluation was conducted on a single dataset, which may restrict generalizability.

Future research may address these limitations by incorporating cost-sensitive or resampling strategies to enhance minority-class detection performance. Hybrid detection architectures that combine unsupervised anomaly ranking with supervised classification may improve discriminative capability. Furthermore, systematic hyperparameter optimization and cross-dataset validation would strengthen empirical robustness and support broader applicability in real-world intrusion detection systems.

The root cause of this poor performance can be attributed to the intrinsic vulnerability of the algorithm when confronted with extreme class imbalance and rigid thresholding. By defining the anomaly threshold based strictly on the contamination parameter (approximately 12.65%), the model was forced to draw a hard decision boundary on the anomaly scores. Because the data is highly imbalanced, this static threshold inadvertently captured a massive amount of normal traffic (False Positives) while leaving true attacks with lower anomaly scores undetected (False Negatives).

This critical failure directly ties back to the primary objective of this study: demonstrating the deceptive nature of traditional evaluation metrics. The stark discrepancy between the ROC-AUC (0.6302) and the PR-AUC (0.1710) serves as quantitative proof that ROC-AUC drastically overestimates model capabilities under highly imbalanced conditions. The large volume of True Negatives artificially inflates the ROC curve, masking the model's critical failure on the minority class. Therefore, this study establishes Isolation Forest not as a standalone solution, but rather as an unsupervised baseline model that highlights the exact vulnerabilities—namely threshold sensitivity and overlap bias—that future hybrid or deep learning models must overcome.

IV. CONCLUSION

This study evaluated the performance of the Isolation Forest algorithm on the UNSW-NB15 dataset under naturally imbalanced class distribution conditions. The experimental results conclusively demonstrate that while the model achieves seemingly high overall accuracy and a moderate ROC-AUC, its operational effectiveness in detecting minority attack instances is critically flawed. The severe failure in attack recall and the consequential high False Negative Rate prove that a substantial portion of malicious traffic bypasses the unsupervised isolation mechanism. Furthermore, the stark discrepancy between precision-recall performance and traditional metrics reveals that accuracy and ROC-AUC provide an overly optimistic and deceptive interpretation of model capability in imbalanced intrusion detection scenarios. From a methodological standpoint, the results highlight the structural limitations of unsupervised anomaly detection models when feature overlap exists between normal and attack traffic. While Isolation Forest offers practical advantages, including scalability and independence from labeled data, it should not be deployed as a standalone intrusion detection mechanism in highly imbalanced environments. In conclusion, this study provides a realistic and quantitatively grounded assessment of Isolation Forest for intrusion detection under class imbalance. The findings underscore the necessity of integrating unsupervised anomaly detection with complementary strategies, adaptive thresholds, or hybrid architectures to enhance detection sensitivity in modern IDS implementation.

ACKNOWLEDGEMENT

This study funded by Institute for Research and Community Service (LPPM) Universitas Pamulang.

REFERENCES

[1] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–22, 2009, doi: 10.1145/1541880.1541882.
[2] A. L. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 2, pp. 1153–1176, Apr. 2016, doi: 10.1109/COMST.2015.2494502.
[3] F. T. Liu, K. M. Ting, and Z. H. Zhou, "Isolation forest," in *Proceedings - IEEE International Conference on Data Mining, ICDM*, 2008, pp. 413–422. doi: 10.1109/ICDM.2008.17.
[4] F. T. Liu, K. M. Ting, and Z. H. Zhou, "Isolation-based anomaly detection," *ACM Trans. Knowl. Discov. Data*, vol. 6, no. 1, Mar. 2012, doi: 10.1145/2133360.2133363.
[5] S. Hariri, M. C. Kind, and R. J. Brunner, "Extended Isolation Forest," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 4, pp. 1479–1489, Apr. 2021, doi: 10.1109/TKDE.2019.2947676.
[6] N. Moustafa and J. Slay, "UNSW-NB15: A Comprehensive Data set for Network Intrusion Detection systems (UNSW-NB15 Network Data Set)." [Online]. Available: https://cve.mitre.org/
[7] S. Meftah, T. Rachidi, and N. Assem, "Network based intrusion detection using the UNSW-NB15 dataset," *International Journal of Computing and Digital Systems*, vol. 8, no. 5, pp. 477–487, 2019, doi: 10.12785/ijcds/080505.
[8] S. M. Kasongo and Y. Sun, "A deep learning method with wrapper based feature extraction for wireless intrusion detection system," *Comput. Secur.*, vol. 92, May 2020, doi: 10.1016/j.cose.2020.101752.
[9] Z. Zoghi and G. Serpen, "Building an intrusion detection system on UNSW-NB15: Reducing the margin of error to deal with data overlap and imbalance," *Concurr. Comput.*, vol. 36, no. 25, Nov. 2024, doi: 10.1002/cpe.8242.
[10] M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, "Building an intrusion detection system using a filter-based feature selection algorithm," *IEEE Transactions on Computers*, vol. 65, no. 10, pp. 2986–2998, Oct. 2016, doi: 10.1109/TC.2016.2519914.
[11] N. Kumar and S. Sharma, "A Hybrid Modified Deep Learning Architecture for Intrusion Detection System with Optimal Feature Selection," *Electronics (Switzerland)*, vol. 12, no. 19, Oct. 2023, doi: 10.3390/electronics12194050.
[12] T. Kieu, B. Yang, C. Guo, and C. S. Jensen, "Outlier Detection for Time Series with Recurrent Autoencoder Ensembles."
[13] N. Sharma, N. S. Yadav, and S. Sharma, "Classification of UNSW-NB15 dataset using Exploratory Data Analysis using Ensemble Learning," *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol. 8, no. 29, 2021, doi: 10.4108/EAI.13-10-2021.171319.
[14] "Mahmud, Md. Saif, et al, "Enhancing Industrial Control System Security An Isolation Forest-based Anomaly Detection Model for Mitigating Cyber Threats", *Journal of Engineering Research and Reports*., vol.26, no.3, 2024, doi: 10.9734/jerr/2024/v26i31102.
[15] X. Tao, Y. Peng, F. Zhao, P. Zhao, and Y. Wang, "A parallel algorithm for network traffic anomaly detection based on Isolation Forest," *Int. J. Distrib. Sens. Netw.*, vol. 14, no. 11, Nov. 2018, doi: 10.1177/1550147718814471.
[16] S. Khalifa, M. Marie, and W. Mohamed, "An Optimized Deep Learning Approach for Multiclass Anomaly Detection," *Information (Switzerland)*, vol. 17, no. 2, Feb. 2026, doi: 10.3390/info17020183.
[17] X. Nong, K. Qin, and X. Xie, "Anomaly Detection in Imbalanced Network Traffic Using a ResCAE-BiGRU Framework," *Symmetry (Basel).*, vol. 17, no. 12, Dec. 2025, doi: 10.3390/sym17122087.
[18] D. Paolini, P. Dini, E. Soldaini, and S. Saponara, "One-Class Anomaly Detection for Industrial Applications: A Comparative Survey and Experimental Study," *Computers*, vol. 14, no. 7, Jul. 2025, doi: 10.3390/computers14070281.
[19] Nalini, M., Yamini, B., Ambhika, C. et al. Enhancing early attack detection: novel hybrid density-based isolation forest for improved anomaly detection. Int. J. Mach. Learn. & Cyber. 16, 3429–3447 (2025). https://doi.org/10.1007/s13042-024-02460-5
[20] P. Verma, D. O'Shea, T. Newe, N. Mehta, N. Bharot, and J. G. Breslin, "ABIDS-VEM: leveraging an equilibrium optimizer and data ramification in association with ensemble learning for anomaly-based intrusion detection system," *Journal of Supercomputing*, vol. 81, no. 7, May 2025, doi: 10.1007/s11227-025-07292-w.
[21] S. Viharika and N. A. Balaji, "Optimized intrusion detection system framework using data sampling, feature selection, and anomaly detection techniques," International Journal of Security and Informatics (IJoSI), vol. 9, no. 2, 2025. doi: 10.6977/IJoSI.202504_9(2).0007.
[22] K. Vani and S. P. Swornambiga, "Adaptive intrusion detection framework for enhanced cloud security in fog and edge computing environments," *International Journal of Advanced Technology and Engineering Exploration*, vol. 11, no. 121, pp. 1613–1640, Dec. 2024, doi: 10.19101/IJATEE.2024.111100395.
[23] E. Ntizikira, L. Wang, J. Chen, and K. Saleem, "Honey-block: Edge assisted ensemble learning model for intrusion detection and prevention using defense mechanism in IoT," Computer Communications, vol. 214, pp. 1–17, 2024, doi: 10.1016/j.comcom.2023.11.023.
[24] M. Mohy-Eddine, A. Guezzaz, S. Benkirane, M. Azrour, and Y. Farhaoui, "An Ensemble Learning Based Intrusion Detection Model for Industrial IoT Security," *Big Data Mining and Analytics*, vol. 6, no. 3, pp. 273–287, Sep. 2023, doi: 10.26599/BDMA.2022.9020032.
[25] A. K. Shukla, S. Srivastav, S. Kumar, and P. K. Muhuri, "UInDeSI4.0: An efficient Unsupervised Intrusion Detection System for network traffic flow in Industry 4.0 ecosystem," *Eng. Appl. Artif. Intell.*, vol. 120, Apr. 2023, doi: 10.1016/j.engappai.2023.105848.
[26] AbuAlghanam, O., Alazzam, H., Alhenawi, E. et al. Fusion-based anomaly detection system using modified isolation forest for internet of things. J Ambient Intell Human Comput 14, 131–145 (2023). https://doi.org/10.1007/s12652-022-04393-9
[27] L. Qi, Y. Yang, X. Zhou, W. Rafique, and J. Ma, "Fast Anomaly Identification Based on Multiaspect Data Streams for Intelligent Intrusion Detection Toward Secure Industry 4.0," *IEEE Trans. Industr. Inform.*, vol. 18, no. 9, pp. 6503–6511, Sep. 2022, doi: 10.1109/TII.2021.3139363.
[28] M. A. N. Anargya et al., "Optimizing IoV Attack Detection using Random Under Sampling Techniques," Jurnal Informatika: Jurnal Pengembangan IT, vol. 10, no. 1, pp. 11–19, 2025, doi: 10.30591/jpit.v10i1.8034.