Deteksi Serangan Siber Jaringan Internet of Things Menggunakan Random Forest pada Dataset IoT-23

Ferri Ojak Immanuel Pardede[1], Parasian D.P. Silitonga[2]
[1]*Program Studi Ilmu Komputer, Universitas HKBP Nommensen Pematangsiantar, Indonesia,*
[2]*Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Katolik Santo Thomas, Medan, Indonesia.*

ARTICLE INFORMATION

Received: April 04, 26
Revised: April 11, 26
Available online: April 30, 26

KEYWORDS

Internet of Things,
Random Forest,
Deteksi Serangan Siber,
Dataset IoT-23,
Machine Learning.

CORRESPONDENCE

Phone: +62 812-6013-3330
E-mail: ferripardede@gmail.com

A B S T R A C T

Internet of Things networks are vulnerable to malicious traffic because connected devices often operate with limited resources and weak security configuration. This study proposes a binary cyberattack detection model for Internet of Things traffic using Random Forest and a labeled IoT-23 connection log subset. The research used a quantitative experimental method with the uploaded conn4_log_labeled.csv file. After metadata cleaning, the dataset contained 156,103 valid records, consisting of 4,536 benign records and 151,567 malicious records. Identifier attributes, detailed labels, timestamps, and raw Internet Protocol addresses were removed to reduce data leakage. Numerical features were converted and imputed using the median, while categorical features were encoded through one-hot encoding. The dataset was split into 80 percent training data and 20 percent testing data using stratified sampling. Random Forest used 100 trees, Gini criterion, balanced class weighting, and random_state 42. The model achieved 99.9968 percent accuracy, 99.9968 percent weighted precision, 99.9968 percent weighted recall, 99.9968 percent weighted F1-score, and 99.9999 percent ROC-AUC. The confusion matrix showed 907 true benign records, 30,313 true malicious records, zero false positives, and one false negative. The most influential features were history and conn_state. These findings show that Random Forest provides a strong and interpretable baseline for binary Internet of Things intrusion detection.

PENDAHULUAN

Internet of Things menghubungkan perangkat fisik, sensor, perangkat lunak, dan jaringan sehingga perangkat dapat mengumpulkan serta bertukar data. Pemakaian perangkat terhubung pada rumah pintar, kamera pengawas, perangkat kesehatan, dan industri memberi manfaat operasional. Namun, konektivitas tersebut juga memperluas permukaan serangan karena perangkat sering memakai konfigurasi bawaan, memiliki sumber daya terbatas, dan tidak selalu mendapat pembaruan keamanan.

Serangan pada lingkungan Internet of Things dapat muncul dalam bentuk pemindaian port, brute force, botnet, command and control, denial of service, dan komunikasi malware. Kondisi ini menuntut intrusion detection system yang mampu mengenali pola trafik berbahaya dari data jaringan. Machine learning relevan karena dapat mempelajari pola dari log koneksi berlabel dan membedakan trafik benign serta malicious berdasarkan fitur aliran jaringan.

IoT-23 menyediakan lalu lintas jaringan Internet of Things berlabel, terdiri atas skenario malware dan benign yang direkam oleh Stratosphere Laboratory, Czech Technical University [1]. Sejumlah penelitian terdahulu memakai dataset ini untuk mengevaluasi model deteksi intrusi berbasis machine learning [5], [6]. Random Forest dipilih karena algoritma ini membangun banyak pohon keputusan dan menggabungkan prediksi melalui voting mayoritas. Mekanisme ensemble tersebut membantu mengurangi overfitting dan tetap menyediakan feature importance untuk interpretasi hasil [3], [4].

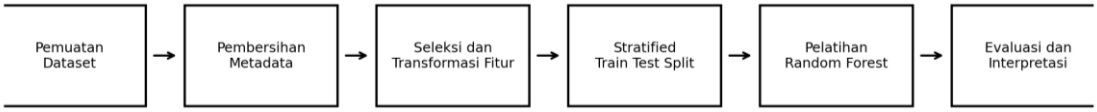
Kebaruan penelitian ini terletak pada penerapan Random Forest pada subset conn4_log_labeled.csv dengan pembersihan atribut pengenal, penggunaan stratified sampling, class_weight balanced, dan pembahasan feature importance pada level fitur utama. Tujuan penelitian ini adalah menguji kinerja Random Forest untuk deteksi serangan siber biner pada subset IoT-23 serta menjelaskan fitur yang paling berpengaruh terhadap keputusan klasifikasi.

Kontribusi utama penelitian ini meliputi tiga poin. Pertama, penelitian menerapkan Random Forest pada subset conn4_log_labeled.csv dari IoT-23 dengan pembersihan atribut yang dirancang untuk mengurangi risiko identifier leakage. Kedua, penelitian menyajikan alur eksperimen yang lengkap, mulai dari capture proses pada bagian metode hingga evaluasi visual melalui confusion matrix, classification report, ROC curve, dan feature importance. Ketiga, penelitian memberikan baseline empiris yang mudah direplikasi untuk deteksi serangan siber biner pada jaringan Internet of Things menggunakan dataset log koneksi berlabel.

METODE PENELITIAN

Penelitian ini menggunakan metode kuantitatif eksperimental berbasis dataset. Eksperimen dilakukan tanpa perakitan alat Internet of Things karena seluruh proses dilakukan melalui komputasi data. Objek penelitian adalah log koneksi berlabel yang telah diubah ke format CSV. Unit analisis berupa satu record koneksi jaringan dengan label Benign atau Malicious.

Tahapan penelitian mencakup pemuatan dataset, pembersihan metadata, seleksi fitur, transformasi label, prapemrosesan fitur, pembagian data latih dan data uji, pelatihan Random Forest, evaluasi model, dan interpretasi feature importance. Alur penelitian ditampilkan pada Gambar 1.



Gambar 1. Alur penelitian deteksi serangan siber berbasis Random Forest

Tabel 1. Ringkasan dataset penelitian

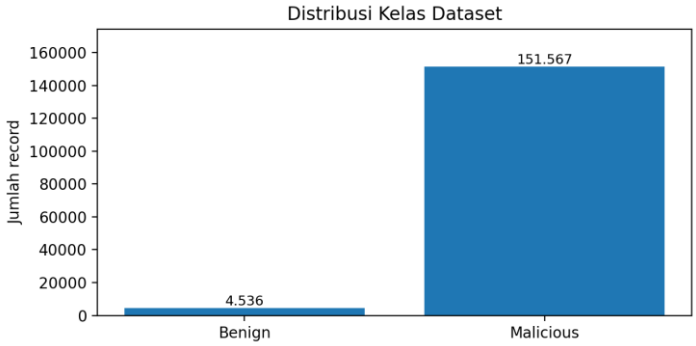
Komponen	Nilai
Nama file	conn4_log_labeled.csv
Jumlah data valid	156.103 record
Kelas benign	4.536 record
Kelas malicious	151.567 record

Data awal berisi 156.104 baris. Satu baris metadata penutup dihapus karena tidak memiliki label kelas. Setelah pembersihan, data valid berjumlah 156.103 record. Distribusi data tidak seimbang karena kelas malicious mencapai 151.567 record, sedangkan kelas benign berjumlah 4.536 record. Ketidakseimbangan ini ditangani dengan stratified split dan class_weight balanced pada model.

Tabel 2. Fitur dan konfigurasi eksperimen

Aspek	Keterangan
Fitur numerik	port, duration, bytes, packets, ip_bytes
Fitur kategorikal	proto, service, conn_state, history
Fitur dihapus	ts, uid, alamat IP mentah, detailed_label, extra
Pembagian data	80% data latih dan 20% data uji
Parameter model	100 pohon, criterion Gini, class_weight balanced

Nilai numerik dikonversi dari teks ke angka dan missing value diimputasi dengan median. Fitur kategorikal dikodekan melalui one-hot encoding. Label target dikonversi menjadi dua kelas, yaitu Benign sebagai 0 dan Malicious sebagai 1. Evaluasi dilakukan memakai accuracy, precision, recall, F1-score, ROC-AUC, dan confusion matrix.



Gambar 2. Distribusi kelas pada dataset penelitian

Untuk memperjelas alur proses eksperimen, Gambar 3 dan Gambar 4 menampilkan capture tahapan pemuatan dataset, pembersihan data, pembagian data latih dan data uji, serta proses pelatihan dan evaluasi model Random Forest. Dokumentasi ini ditempatkan pada bagian metode agar urutan eksperimen lebih mudah diikuti sebelum pembaca masuk ke bagian hasil dan pembahasan.

```
Capture 1 - Pemuatan dataset dan pembagian data
>>> import pandas as pd
>>> df = pd.read_csv("conn4_log_labeled.csv", header=None, skiprows=1)
>>> df = df[df["label"].isin(["Benign", "Malicious"])]

# Hasil pemeriksaan dataset
Jumlah data valid        : 156.103 record
Jumlah fitur awal        : 24 kolom
Kelas Malicious          : 151.567 record
Kelas Benign             :   4.536 record

>>> train_test_split(test_size=0.2, stratify=y, random_state=42)
Data latih               : 124.882 record
Data uji                 :  31.221 record
```

Gambar 3. Capture proses pemuatan dataset dan pembagian data

Capture 2 - Pelatihan dan evaluasi Random Forest

```
>>> RandomForestClassifier(n_estimators=100, criterion="gini",
...     class_weight="balanced", n_jobs=-1, random_state=42)
>>> pipeline.fit(X_train, y_train)
>>> y_pred = pipeline.predict(X_test)

# Hasil evaluasi model
Accuracy             : 0,999968  atau 99,9968%
Weighted precision   : 0,999968  atau 99,9968%
Weighted recall      : 0,999968  atau 99,9968%
Weighted F1-score    : 0,999968  atau 99,9968%
ROC-AUC              : 0,99999998

Confusion Matrix:
[[  907     0]
 [    1 30313]]
```

Gambar 4. Proses pelatihan dan evaluasi Random Forest

(1)

HASIL DAN PEMBAHASAN

Pembagian data dengan stratified sampling menjaga komposisi kelas pada data latih dan data uji. Data latih berjumlah 124.882 record, terdiri dari 121.253 malicious dan 3.629 benign. Data uji berjumlah 31.221 record, terdiri dari 30.314 malicious dan 907 benign.

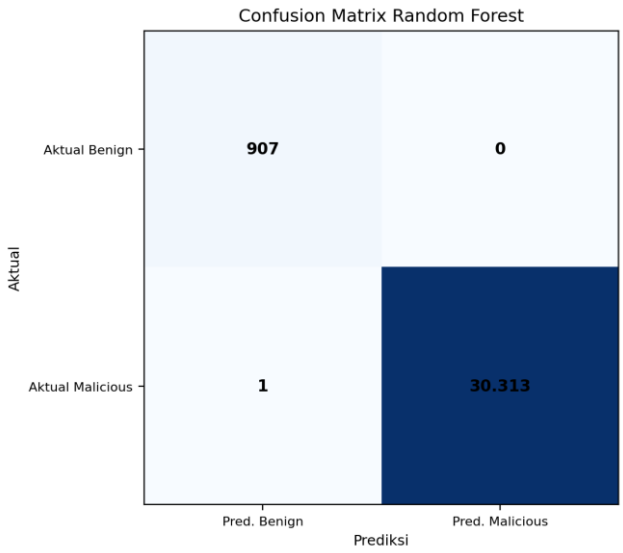
Tabel 3. Distribusi data latih dan data uji

Subset	Benign	Malicious	Total
Data latih	3.629	121.253	124.882
Data uji	907	30.314	31.221
Total	4.536	151.567	156.103

Random Forest memperoleh performa sangat tinggi pada data uji. Nilai accuracy, weighted precision, weighted recall, dan weighted F1-score sama-sama mencapai 99,9968 persen. Nilai ROC-AUC mencapai 99,9999 persen. Hasil ini menunjukkan kemampuan separasi kelas yang sangat kuat pada subset yang diuji.

Tabel 4. Hasil evaluasi model

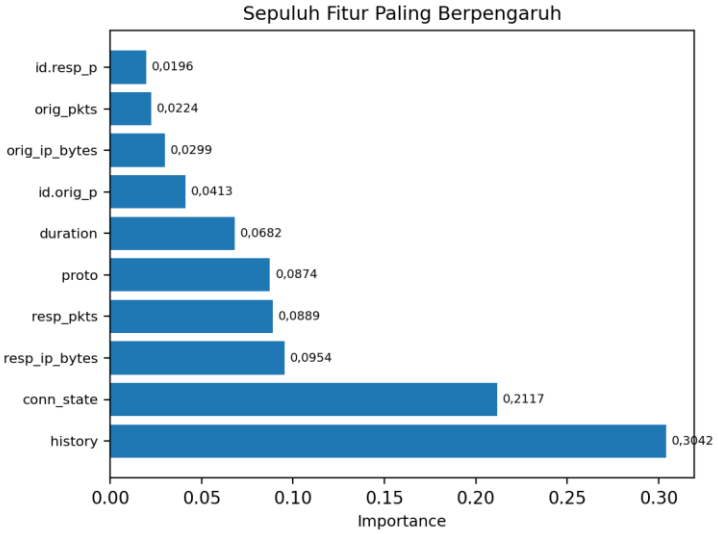
Metrik	Nilai	Persentase
Accuracy	0,999968	99,9968%
Precision tertimbang	0,999968	99,9968%
Recall tertimbang	0,999968	99,9968%
F1-score tertimbang	0,999968	99,9968%
ROC-AUC	0,99999998	99,9999%



Gambar 5. Confusion matrix hasil pengujian model

Confusion matrix menunjukkan 907 true benign, 30.313 true malicious, nol false positive, dan satu false negative. Tidak adanya false positive berarti seluruh trafik benign pada data uji tidak salah ditandai sebagai serangan. Satu false negative perlu mendapat perhatian karena trafik malicious yang lolos dapat menurunkan efektivitas deteksi pada sistem nyata.

Hasil per kelas memperlihatkan bahwa model tetap stabil pada kelas minoritas. Precision benign mencapai 0,998899, recall benign mencapai 1,000000, dan F1-score mencapai 0,999449. Pada kelas malicious, precision mencapai 1,000000, recall mencapai 0,999967, dan F1-score mencapai 0,999984. Temuan ini menunjukkan bahwa penggunaan stratified split dan class_weight balanced membantu mengurangi bias terhadap kelas mayoritas.



Gambar 6. Feature importance pada model Random Forest

Tabel 5. Sepuluh fitur paling berpengaruh

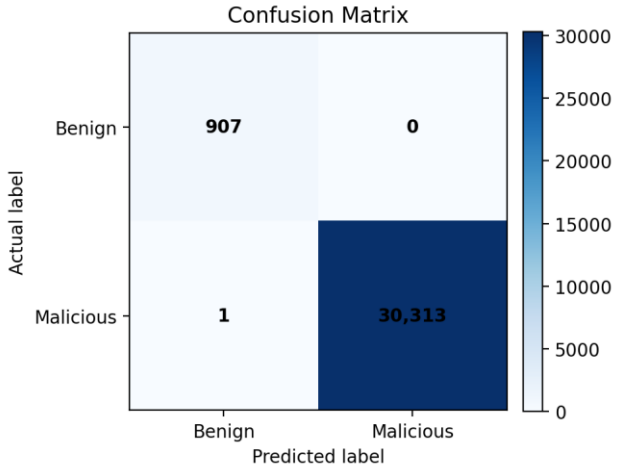
No	Fitur	Importance
1	history	0,304214
2	conn_state	0,211672
3	resp_ip_bytes	0,095395
4	resp_pkts	0,088945
5	proto	0,087417
6	duration	0,068190
7	id.orig_p	0,041296
8	orig_ip_bytes	0,029865
9	orig_pkts	0,022366
10	id.resp_p	0,019563

Feature importance menunjukkan bahwa history dan conn_state menjadi dua atribut paling dominan. Kedua fitur tersebut berkaitan dengan riwayat komunikasi dan status koneksi, sehingga wajar jika berperan besar dalam membedakan trafik normal dan berbahaya. Fitur resp_ip_bytes, resp_pkts, proto, dan duration juga memberi kontribusi penting karena menggambarkan pola respons dan karakteristik durasi komunikasi.

Temuan penelitian ini mendukung hasil studi sebelumnya yang menunjukkan bahwa metode ensemble efektif untuk deteksi intrusi Internet of Things [5], [8], [9]. Keunggulan Random Forest terletak pada akurasi tinggi, kemudahan penerapan, dan kemampuan memberikan interpretasi fitur. Namun, hasil yang sangat tinggi harus dibaca secara hati-hati karena dataset berasal dari satu subset log koneksi dan distribusi kelas sangat tidak seimbang.

Secara praktis, model dapat dijadikan baseline awal untuk intrusion detection system pada jaringan Internet of Things. Pada implementasi nyata, sistem dapat mengekstrak fitur koneksi dengan Zeek atau alat sejenis, menjalankan pipeline prapemrosesan yang sama, lalu mengirim fitur ke model Random Forest. Sebelum digunakan pada jaringan operasional, model perlu divalidasi ulang dengan data lokal dan beberapa skenario serangan.

Sebagai penguatan hasil pembahasan, evaluasi model juga ditampilkan dalam bentuk visual. Gambar 7 menyajikan confusion matrix yang lebih visual, Gambar 8 menampilkan classification report, dan Gambar 9 menampilkan kurva ROC. Ketiga visual ini membantu pembaca melihat performa model secara ringkas dan mendukung interpretasi hasil numerik yang telah disajikan pada tabel.

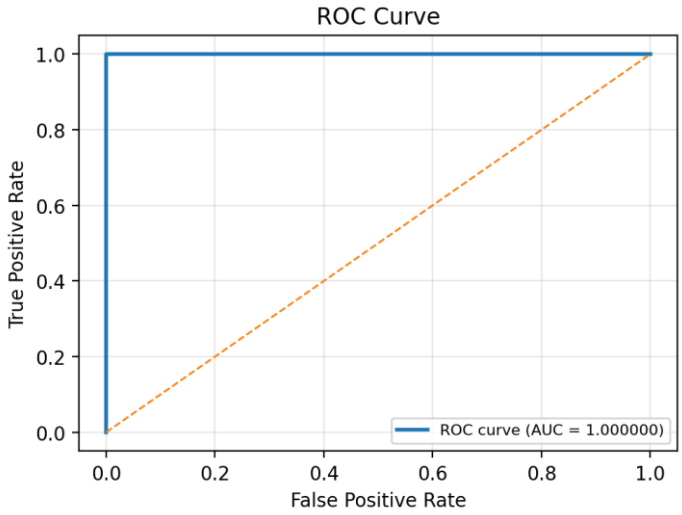


Gambar 7. Diagram confusion matrix yang lebih visual

Classification Report

	precision	recall	f1-score	support
Benign	0.998899	1.000000	0.999449	907
Malicious	1.000000	0.999967	0.999984	30,314
accuracy			0.999968	31,221
macro avg	0.999449	0.999984	0.999716	31,221
weighted avg	0.999968	0.999968	0.999968	31,221

Gambar 8. Visual classification report hasil pengujian model



Gambar 9. Kurva ROC model Random Forest

Dari sisi validitas eksperimen, penelitian ini tidak memakai uid, id.orig_h, id.resp_h, detailed_label, dan atribut identitas lain sebagai fitur. Keputusan ini penting karena atribut pengenal dapat membuat model menghafal capture tertentu, bukan mempelajari pola serangan. Dengan demikian, fitur yang dipertahankan lebih berorientasi pada karakteristik aliran koneksi, seperti port, protokol, durasi, jumlah byte, jumlah paket, status koneksi, dan riwayat komunikasi.

Walaupun model menghasilkan nilai evaluasi yang sangat tinggi, interpretasi hasil tetap perlu mempertimbangkan ketidakseimbangan kelas. Accuracy tinggi dapat terlihat meyakinkan, tetapi precision, recall, F1-score, ROC-AUC, dan confusion matrix tetap harus dibaca bersama. Pada penelitian ini, seluruh metrik tersebut menunjukkan konsistensi hasil. Namun, penggunaan satu subset log koneksi membuat hasil lebih tepat diposisikan sebagai baseline empiris, bukan klaim final untuk seluruh skenario IoT-23.

Dari sisi penerapan, pipeline yang dibangun dapat dikembangkan menjadi rancangan awal network intrusion detection system. Pada sistem nyata, log koneksi dapat diekstraksi dengan Zeek, diproses melalui skema prapemrosesan yang sama, lalu diklasifikasikan oleh model Random Forest. Output model dapat digunakan sebagai peringatan awal

bagi administrator. Implementasi tersebut tetap membutuhkan validasi ulang memakai data lokal karena pola trafik setiap jaringan Internet of Things dapat berbeda.

KESIMPULAN

Penelitian ini berhasil membangun model deteksi serangan siber pada jaringan Internet of Things menggunakan Random Forest berbasis subset Dataset IoT-23. Dataset valid berisi 156.103 record dengan 151.567 malicious dan 4.536 benign. Data dibagi menjadi 124.882 data latih dan 31.221 data uji melalui stratified sampling.

Model memperoleh accuracy 99,9968 persen, weighted precision 99,9968 persen, weighted recall 99,9968 persen, weighted F1-score 99,9968 persen, dan ROC-AUC 99,9999 persen. Confusion matrix menunjukkan 907 true benign, 30.313 true malicious, nol false positive, dan satu false negative. Fitur paling penting adalah history, conn_state, resp_ip_bytes, resp_pkts, proto, dan duration.

Keterbatasan penelitian ini terletak pada penggunaan satu file subset dan ketidakseimbangan kelas. Penelitian lanjutan perlu menggunakan beberapa capture IoT-23, melakukan validasi lintas skenario, menerapkan hyperparameter tuning, dan membandingkan Random Forest dengan algoritma ensemble lain seperti XGBoost atau LightGBM.

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada penyedia Dataset IoT-23 dari Stratosphere Laboratory, Czech Technical University, serta semua pihak yang mendukung proses penyusunan artikel ini.

REFERENSI

[1] S. Garcia, A. Parmisano, and M. J. Erquiaga, “IoT-23: A labeled dataset with malicious and benign IoT network traffic,” Zenodo, 2020, doi: 10.5281/zenodo.4743746.
[2] J. Vitorino, R. Andrade, I. Praça, O. Sousa, and E. Maia, “A comparative analysis of machine learning techniques for IoT intrusion detection,” in Foundations and Practice of Security, Lecture Notes in Computer Science, vol. 13291, Cham: Springer, 2022, pp. 191-207, doi: 10.1007/978-3-031-08147-7_13.
[3] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5-32, 2001, doi: 10.1023/A:1010933404324.
[4] F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.
[5] S. Strecker, R. Dave, N. Siddiqui, and N. Seliya, “A modern analysis of aging machine learning based IoT cybersecurity methods,” arXiv preprint arXiv:2110.07832, 2021.
[6] M. A. Akif, I. Butun, A. Williams, and I. Mahgoub, “Hybrid machine learning models for intrusion detection in IoT: Leveraging a real-world IoT dataset,” arXiv preprint arXiv:2502.12382, 2025.
[7] E. Krzysztoń, P. Rumiński, and M. Domański, “A comparative analysis of anomaly detection methods in network traffic,” Applied Sciences, vol. 14, no. 24, p. 11545, 2024.
[8] G. Airlangga, “Comparative analysis of machine learning models for intrusion detection system in IoT networks,” MALCOM: Indonesian Journal of Machine Learning and Computer Science, 2024.
[9] B. M. Kouassi, A. A. Abdou, and M. A. Ferrag, “Top-K feature selection for IoT intrusion detection,” Future Internet, vol. 17, no. 11, p. 529, 2025.
[10] M. Nicho, A. Alazab, and M. Alazab, “Assessing IoT intrusion detection computational costs,” Journal of Cyber Security Technology, 2025.
[6] J. Vitorino, I. Praça, and E. Maia, “Towards adversarial realism and robust learning for IoT intrusion detection and classification,” arXiv preprint arXiv:2301.13122, 2023.
[7] K. G. R. Narayan, S. Mookherji, V. Odelu, R. Prasath, A. C. Turlapaty, and A. K. Das, “IIDS: Design of intelligent intrusion detection system for Internet-of-Things applications,” arXiv preprint arXiv:2308.00943, 2023.
[8] G. Airlangga, “Comparative analysis of machine learning models for intrusion detection in Internet of Things networks using the RT-IoT2022 dataset,” MALCOM: Indonesian Journal of Machine Learning and Computer Science, 2024.
[9] G. Guntoro, “Optimizing Random Forest for IoT cyberattack detection with imbalance handling,” MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer, 2025.
[10] B. M. Kouassi, A. A. Abdou, and M. A. Ferrag, “Top-K feature selection for IoT intrusion detection,” Future Internet, vol. 17, no. 11, p. 529, 2025.