Jurnal Informatika
E-ISSN: 2528-0090
P-ISSN: 2252-9160

Vol. 15, No. 1, Juni 2026, Page. 43~59
https://ejournal.unidayan.ac.id/index.php/JIU
DOI: https://doi.org/10.55340/jiu.v15i1.2685

Seleksi Fitur dan Optimasi Model untuk Sistem Deteksi Intrusi Berbasis Machine Learning pada Jaringan IoT: Sebuah Tinjauan Literatur Sistematis

Feature Selection and Model Optimization for Machine Learning-Based Intrusion Detection Systems in IoT Networks: A Systematic Literature Review

Abdul Burhanudin*[1)], Adhwa Pranaja Widyadana [2)], Arip Solehudin[3)]
[1,2,3] Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang
Email: *[1] 2310631170001@student.unsika.ac.id

Article Info:	Received	Revised	Accepted	Published
	13 Mei 2026	13 Mei 2026	03 Juni 2026	03 Juni 2026

Abstrak

Pertumbuhan pesat jaringan Internet of Things (IoT) telah memperluas permukaan serangan siber, sehingga Intrusion Detection System (IDS) yang akurat dan efisien menjadi semakin diperlukan. IDS berbasis machine learning telah banyak diadopsi untuk mengidentifikasi pola lalu lintas jaringan normal dan berbahaya, namun kinerjanya sangat dipengaruhi oleh kualitas fitur, karakteristik dataset, dan strategi optimasi model. Penelitian ini bertujuan menganalisis pengaruh metode seleksi fitur dan optimasi model terhadap kinerja IDS berbasis machine learning pada jaringan IoT. Systematic Literature Review dilakukan terhadap 25 artikel jurnal nasional dan internasional yang diterbitkan antara tahun 2020 hingga 2025, dengan analisis melalui tahap identifikasi, seleksi, ekstraksi data, dan sintesis tematik. Hasil kajian menunjukkan bahwa seleksi fitur mampu mengurangi atribut yang tidak relevan, menurunkan kompleksitas komputasi, dan meningkatkan stabilitas klasifikasi. Optimasi model melalui Bayesian Optimization, Genetic Algorithm, SMOTE, normalisasi, dan ekstraksi fitur turut mendukung peningkatan kinerja IDS. Temuan ini menegaskan bahwa kinerja IDS tidak cukup dinilai hanya dari akurasi, tetapi juga harus mempertimbangkan precision, recall, F1-score, false alarm rate, jumlah fitur terpilih, dan waktu komputasi. Penelitian ini berkontribusi secara teoretis dengan memetakan hubungan antara seleksi fitur, optimasi model, dan kinerja IDS pada jaringan IoT, sekaligus memberikan panduan praktis bagi pengembangan sistem deteksi intrusi yang lebih adaptif.

Kata Kunci: *Deteksi Intrusi, Internet of Things, Machine Learning, Optimasi Model.*

Jurnal Informatika
E-ISSN: 2528-0090
P-ISSN: 2252-9160
Vol. 15, No. 1, Juni 2026, Page. 43~59
https://ejournal.unidayan.ac.id/index.php/JIU
DOI: https://doi.org/10.55340/jiu.v15i1.2685

Abstract

The rapid growth of Internet of Things (IoT) networks has expanded the cyberattack surface, making accurate and efficient Intrusion Detection Systems (IDS) increasingly necessary. Machine learning-based IDS has been widely adopted for identifying patterns of normal and malicious network traffic, yet its performance remains strongly influenced by feature quality, dataset characteristics, and model optimization strategies. This study analyzes the effect of feature selection methods and model optimization on the performance of machine learning-based IDS in IoT networks. A Systematic Literature Review was conducted on 25 relevant Indonesian and international journal articles published between 2020 and 2025, with analysis performed through identification, selection, data extraction, and thematic synthesis. The findings indicate that feature selection reduces irrelevant attributes, lowers computational complexity, and improves classification stability. Model optimization through Bayesian Optimization, Genetic Algorithm, SMOTE, normalization, and feature extraction further supports IDS performance improvement. These findings confirm that IDS performance should be assessed not only through accuracy, but also through precision, recall, F1-score, false alarm rate, selected feature count, and computational time. This study contributes theoretically by mapping the relationship between feature selection, model optimization, and IDS-IoT performance, while also offering practical guidance for developing more adaptive intrusion detection systems.

Keywords: *feature selection, Internet of Things, machine learning, model optimization.*

1. PENDAHULUAN

Pertumbuhan konektivitas global menjadikan jaringan *Internet of Things* (IoT) sebagai infrastruktur utama dalam aktivitas digital modern, seiring dengan meningkatnya ancaman siber yang semakin kompleks dan adaptif dalam skala maupun variasinya [1],[2],[3],[4],[5]. Metode keamanan konvensional dinilai kurang efektif menghadapi serangan seperti DoS, DDoS, *botnet*, *malware*, dan intrusi berbasis anomali, sehingga pendekatan berbasis *machine learning* dan *deep learning* banyak diadopsi karena mampu mengenali pola lalu lintas jaringan secara otomatis dan memberikan deteksi *real-time* dengan akurasi tinggi [6],[7],[8],[9],[10]. Namun, tantangan utama dalam pengembangan IDS berbasis *machine learning* terletak pada tingginya dimensionalitas data jaringan, di mana fitur tidak relevan atau redundan meningkatkan beban komputasi sekaligus menurunkan kemampuan generalisasi model terhadap serangan baru [11],[12],[13],[14],[15]. Ketidakseimbangan distribusi kelas dalam dataset turut menjadi kendala serius karena model cenderung bias terhadap kelas mayoritas dan gagal mendeteksi serangan dari kelas minoritas [16],[17],[18].

Sebagai upaya mengatasi tantangan prapemrosesan tersebut, beberapa penelitian terdahulu telah dilakukan secara parsial. Penelitian yang dilakukan oleh Kikissagbe dan Adda [19] menemukan bahwa pemilihan algoritma klasifikasi secara terisolasi kurang efektif tanpa

prapemrosesan yang kuat, sementara penelitian oleh Morshedi dan Matinkhah [20] menunjukkan bahwa integrasi *deep learning* memerlukan representasi fitur yang tepat agar performa deteksi optimal. Di sisi lain, penelitian Gaspar et al. [21] berfokus pada penerapan *explainable AI* untuk melihat transparansi penentuan keputusan model, sedangkan penelitian Abdulkareem [22] mengusulkan algoritma *metaheuristic* khusus untuk seleksi fitur bernilai tinggi. Selain itu, penelitian Khairullah dan Alsenani [23] serta Elnakib et al. [24] membuktikan bahwa arsitektur *deep learning* menghasilkan performa yang menjanjikan pada perangkat IoT, sementara penelitian Ahsan et al. [25] menegaskan bahwa validasi lintas dataset pada metode hibrida masih menjadi hambatan utama yang belum terselesaikan.

Meskipun penelitian-penelitian tersebut telah memberikan kontribusi penting, terdapat gap penelitian yang mendasar dan belum terjawab secara holistik dalam literatur saat ini. Celah pertama terlihat dari adanya fragmentasi dalam menilai kinerja IDS, di mana sebagian besar studi terdahulu masih mengevaluasi model secara terpisah (hanya berfokus pada algoritma saja atau seleksi fitur saja) tanpa melihat keterpaduannya secara sistematis. Celah kedua adalah belum seragamnya standar pengukuran performa, karena literatur terdahulu masih terjebak pada penggunaan akurasi sebagai metrik tunggal dan kerap mengabaikan metrik operasional yang krusial bagi ekosistem IoT, seperti efisiensi waktu komputasi (*runtime*) dan tingkat alarm palsu (*false alarm rate*). Selain itu, belum ada protokol evaluasi tunggal yang menguji interaksi simultan antara reduksi dimensi data dan optimasi penyeimbangan kelas dalam satu *pipeline* terintegrasi.

Untuk menjembatani celah tersebut, penelitian ini bertujuan menganalisis dan mensintesis secara komprehensif pengaruh metode seleksi fitur dan optimasi model terhadap kinerja IDS berbasis *machine learning* pada jaringan IoT melalui pendekatan *Systematic Literature Review* (SLR). Sebagai panduan utama dalam ekstraksi dan sintesis data secara terstruktur, penelitian ini memformulasikan empat rumusan masalah formal (*Research Questions/RQ*), yaitu: (RQ1) apa saja *dataset benchmark* yang paling relevan dan banyak dikaji dalam penelitian IDS untuk jaringan IoT; (RQ2) pendekatan metode seleksi fitur apa yang paling efektif diterapkan untuk mereduksi karakteristik data IoT berdimensi tinggi; (RQ3) bagaimana tren strategi optimasi model yang dominan diterapkan untuk menangani ketidakseimbangan kelas (*imbalanced data*); serta (RQ4) bagaimana dampak komparatif dari implementasi seleksi fitur dan optimasi model terhadap metrik operasional IDS.

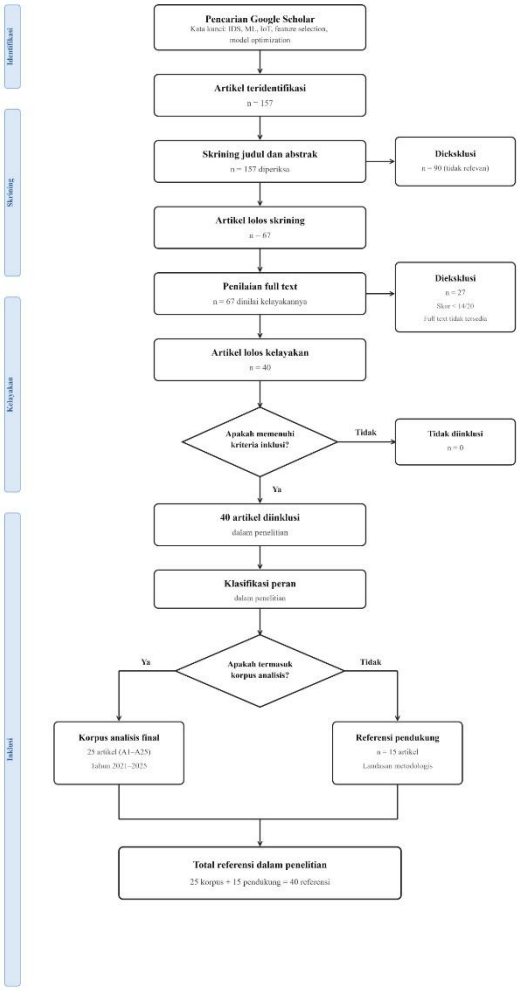
Berbeda dengan SLR sebelumnya seperti kajian Mallidi [1] yang berfokus luas pada optimasi prapemrosesan umum atau survei Momand et al. [2] yang menitikberatkan pada taksonomi serangan, kebaruan dan kontribusi utama dari penelitian ini terletak pada penyusunan sintesis konseptual yang mengonfrontasi interaksi langsung antara reduksi dimensi data dan penanganan ketidakseimbangan kelas terhadap metrik operasional perangkat IoT. Melalui perbandingan komparatif yang berimbang antara literatur nasional terakreditasi dan internasional bereputasi, penelitian ini memberikan kontribusi berupa usulan protokol evaluasi multi-metrik yang dapat dipertanggungjawabkan bagi pengembang sistem keamanan IoT berskala besar di masa depan.

Jurnal Informatika Vol. 15, No. 1, Juni 2026, Page. 43~59
E-ISSN: 2528-0090 https://ejournal.unidayan.ac.id/index.php/JIU
P-ISSN: 2252-9160 DOI: https://doi.org/10.55340/jiu.v15i1.2685

2. METODE

2.1 Desain dan Kerangka SLR (PRISMA)

Penelitian ini menggunakan desain *Systematic Literature Review* (SLR) dengan pendekatan kualitatif deskriptif. Kerangka pelaksanaan penelitian ini berpedoman secara ketat pada protokol *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) yang terdiri dari empat tahapan utama: identifikasi, skrining, penilaian kelayakan, dan inklusi. Tujuan penggunaan kerangka PRISMA adalah untuk memastikan transparansi, konsistensi, dan replikabilitas proses seleksi artikel.



Gambar 1. Diagram alur seleksi artikel

Gambar 1 mengilustrasikan tahapan seleksi artikel berdasarkan protokol PRISMA, mulai dari pencarian awal (157 artikel) hingga klasifikasi peran menggunakan simbol keputusan (decision) untuk memisahkan 25 artikel korpus utama dan 15 artikel referensi pendukung.

2.2 Research Questions (RQ)

Untuk memandu arah ekstraksi dan sintesis data secara sistematis, penelitian ini merumuskan empat *Research Questions* (RQ) utama sebagai berikut:

- RQ1: Apa saja dataset *benchmark* yang paling relevan dan banyak digunakan dalam penelitian IDS untuk jaringan IoT?
- RQ2: Pendekatan dan metode seleksi fitur apa yang paling efektif diterapkan pada karakteristik data IoT berdimensi tinggi?
- RQ3: Bagaimana tren strategi optimasi model yang diterapkan untuk menangani ketidakseimbangan kelas (*imbalanced data*)?
- RQ4: Bagaimana dampak komparatif dari implementasi seleksi fitur dan optimasi model terhadap performa metrik evaluasi?

2.3 Sumber Pencarian & Search String

Proses pencarian awal dilakukan secara komprehensif pada basis data akademik melalui mesin pencari Google Scholar. Untuk memastikan spesifisitas hasil pencarian, *search string* boolean yang digunakan adalah: ("Intrusion Detection System" OR "IDS") AND ("Internet of Things" OR "IoT") AND ("Machine Learning" OR "Deep Learning") AND ("Feature Selection" OR "Dimensionality Reduction") AND "Model Optimization". Pencarian ini dibatasi pada artikel yang terbit dalam rentang lima tahun terakhir (2020–2025).

2.4 Kriteria Inklusi dan Eksklusi

Kriteria inklusi yang ditetapkan mencakup: (1) Artikel jurnal nasional terakreditasi atau internasional bereputasi; (2) Diterbitkan antara tahun 2020 hingga 2025; (3) Fokus pada IDS di jaringan IoT menggunakan *machine learning*; (4) Membahas seleksi fitur dan/atau optimasi model; serta (5) Melaporkan metrik evaluasi kinerja secara eksplisit. Sebaliknya, kriteria eksklusi diterapkan untuk mengeliminasi: karya non-jurnal (skripsi, makalah populer), artikel *review* murni tanpa eksperimen model, serta artikel yang tidak memiliki akses *full text*.

2.5 Penilaian Kualitas Artikel

Penilaian kualitas dilakukan menggunakan 10 pertanyaan diagnostik dengan ambang batas penerimaan skor minimal 14 dari total bobot 20. Instrumen ini mengevaluasi kejelasan tujuan, ketepatan metodologi, penjelasan rincian dataset, validitas rancangan uji coba, dan kelengkapan pelaporan hasil (termasuk metrik komputasi). Artikel yang gagal mencapai skor batas tersebut dieliminasi dari korpus utama.

2.6 Instrumen Ekstraksi Data

Data dari setiap artikel yang lolos ekstraksi dicatat secara terstruktur menggunakan instrumen khusus untuk memastikan konsistensi. Terdapat 14 komponen informasi yang diekstrak dari setiap artikel yang nantinya menjadi dasar analisis sintesis tematik untuk menjawab seluruh *Research Questions* (RQ). Detail komponen ekstraksi disajikan pada Tabel 1.

Tabel 1. Instrumen Ekstraksi Data Artikel

No	Komponen Ekstraksi	Keterangan yang Dicatat
1	Kode Artikel	A1–A25
2	Penulis dan Tahun	Nama penulis utama dan tahun publikasi
3	Judul Artikel	Judul lengkap artikel jurnal
4	Nama Jurnal	Nama jurnal tempat artikel diterbitkan
5	Tujuan Penelitian	Fokus utama penelitian terdahulu
6	Jenis Jaringan	Jaringan komputer, IoT, server, botnet, atau keamanan jaringan
7	Dataset	NSL-KDD, UNSW-NB15, CICIDS, Honeynet BSSN, atau lainnya

8	Metode Seleksi Fitur	Information Gain, Correlation, CSE, atau lainnya
9	Metode Optimasi Model	Bayesian Optimization, SMOTE, tuning parameter, atau lainnya
10	Algoritma ML	Random Forest, Decision Tree, SVM, XGBoost, atau lainnya
11	Metrik Evaluasi	Accuracy, precision, recall, F1-score, FAR, waktu komputasi
12	Hasil Utama	Temuan utama dari artikel
13	Keterbatasan	Kekurangan metode, dataset, metrik, atau validasi
14	Relevansi terhadap SLR	Hubungan dengan topik seleksi fitur, optimasi model, IDS-IoT

Tabel 1 menunjukkan 14 komponen data yang diekstrak dari setiap artikel korpus. Poin-poin ini digunakan sebagai acuan baku untuk mengidentifikasi metodologi, dataset, algoritma, dan temuan secara seragam.

Lembar penilaian kualitas artikel dilakukan menggunakan 10 kriteria diagnostik dengan ambang batas kelulusan skor minimal 14 dari 20. Kriteria tersebut mengevaluasi kesesuaian rentang tahun publikasi, bahasa dan reputasi jurnal, relevansi topik IDS dan IoT, penggunaan *machine learning*, serta kejelasan deskripsi dataset dan metrik evaluasi. Untuk menjaga objektivitas proses seleksi, ditetapkan kriteria inklusi yang mewajibkan artikel berasal dari jurnal nasional terakreditasi atau internasional bereputasi (terbit 2020–2025), secara eksplisit membahas IDS, IoT, *machine learning*, seleksi fitur, optimasi model, serta menyajikan metrik kinerja. Sebaliknya, kriteria eksklusi diterapkan untuk mengeliminasi literatur non-jurnal ilmiah (seperti opini atau skripsi), artikel di luar rentang tahun, artikel tanpa akses *full text*, dan studi yang tidak memuat evaluasi model secara utuh.

Tahap akhir instrumen adalah penyusunan tabel rekap 25 artikel final yang memuat kode artikel, penulis, dataset, metode seleksi fitur, optimasi model, algoritma, metrik evaluasi, temuan utama, dan kolom gap penelitian. Tabel rekap ini menjadi kerangka utama dalam proses sintesis tematik pada bagian hasil dan pembahasan.

3. HASIL

3.1 Rekap Artikel Korpus

Tabel 2 menyajikan rekap 25 artikel korpus (A1–A25) yang menjadi dasar sintesis tematik dalam SLR ini.

Tabel 2. Rekap 25 Artikel Korpus SLR

Kode	Penulis (Tahun)	Dataset	Seleksi Fitur	Optimasi Model	Algoritma	Metrik	Temuan Utama
A1	Bakır & Ceviz [26] (2024)	CICIDS-2017	Hybrid FS	GA hyperparameter	RF, DT	Accuracy, F FAR	GA-based tuning perkuat performa IDS
A2	Agustina et al. [27] (2024)	UNSW-NB15	Filter FS	—	Random Forest	Accuracy, Recall, F1	FS tingkatkan sensitivitas RF pada anomali
A3	Jaw & Wang [28] (2021)	UNSW-NB15	Ensemble-based FS	—	RF, ensemble	Accuracy, Precision, Recall	Ensemble FS efektif pada data dimensi tinggi
A4	Logeswari et al. [29] (2025)	NSL-KDD, UNSW-NB15	Dual-layer FS	—	SVM, RF	Accuracy, F FAR	Dual-layer FS lebih efisien dari FS tunggal
A5	Kurniabudi et al. [30] (2022)	CICIDS-2017	Classifier Subset Evaluator	SMOTE	RF, NB	Accuracy, Precision	CSE optimalkan seleksi fitur IoT
A6	Khoo & Handoko [31] (2025)	Botnet datase	Filter FS	—	RF, DT, KNN	Accuracy, F	Perbandingan ML berbasis FS deteksi botnet

A7	Kunhare et al. [32] (2022)	NSL-KDD	Genetic Algorithm	Genetic Algorithm	SVM, RF, hybrid	Accuracy, F	GA efektif untuk FS dan optimasi classifier
A8	Sayegh et al. [33] (2024)	CIC-IDS2017	Filter FS	SMOTE, LSTM	LSTM	Accuracy, F FAR	SMOTE+LSTM atasi ketidakseimbangan kelas
A9	Hariyanti et al. [34] (2024)	CICIDS-2017	Wrapper FS	—	Decision Tree	Accuracy, F FAR	Perbandingan metode FS pada DT
A10	Kocher & Kumar [35] (2021)	UNSW-NB15	Chi-Square	—	SVM, NB, DT, RF	Accuracy, FAR	Chi-Square reduksi fitur tanpa turun akurasi
A11	Kareem et al. [36] (2022)	NSL-KDD, UNSW-NB15	Hybrid metaheuristic	—	KNN, SVM	Accuracy, FAR	Hybrid metaheuristic ungguli filter tunggal
A12	Putri et al. [37] (2021)	UNSW-NB15	IGR + Correlation	—	Naive Bayes, J48	Accuracy, F	Kombinasi IGR+Correlation diskriminatif
A13	Polatgil [38] (2022)	KDD Cup99	—	Normalization	DT, SVM, RF	Accuracy	Normalisasi pengaruhi performa klasifikasi IDS
A14	Prihantono & Ramli [39] (2022)	CICIDS-2017	Model-based FS	—	RF, XGBoost	Accuracy, F	Model-based FS tingkatkan deteksi serangan
A15	Suryadi & Marzuki [40] (2023)	NSL-KDD	—	—	RF, DT, SVM	Accuracy, Precision	Perbandingan algoritma ML untuk IDS
A16	Yin et al. [41] (2023)	UNSW-NB15	IGRF-RFE hybrid	—	MLP	Accuracy, F FAR	IGRF-RFE lebih optimal dari metode tunggal
A17	Inayah & Ramli [42] (2024)	Honeynet BSSN	—	Class balancing	Random Forest	Accuracy, Precision, Recall, F1	RF pada dataset tidak seimbang
A18	Samsudiat & Ramli [43] (2025)	CICIoT2023	Hybrid FS	Bayesian Optimization	RF, XGBoost	Accuracy, F waktu komputasi	Hybrid FS + Bayesian Opt tingkatkan IDS-IoT
A19	Putra & Amarudin [44] (2025)	NSL-KDD	—	Hyperparameter tuning	RF, SVM, DT	Accuracy, F Recall	Perbandingan ML untuk IDS NSL-KDD
A20	Faizin et al. [45] (2024)	NSL-KDD	Thresholding	—	RF, DT	Accuracy, F	Thresholding adaptif optimalkan batas FS
A21	Li et al. [46] (2024)	TON-IoT	FS vs Feature Extraction	—	RF, XGBoost	Accuracy, F waktu komputasi	FS dan FE perlu dibandingkan dalam konteks sama
A22	Shiddiq et al. [47] (2025)	CICIDS-2017	FS + Feature Extraction	Oversampling	RF, DT	Accuracy, Precision, F	Kombinasi oversampling+FS+FE konsisten
A23	Walling & Lodh [48] (2024)	NSL-KDD, UNSW-NB15	Hybrid FS statistik	—	SVM, RF	Accuracy, Precision, FAR	Hybrid FS statistik tingkatkan NIDS-IoT
A24	Khashan & Khafajah [49] (2024)	CICIoT2023	ML-based FS + ensemble	—	RF, ensemble	Accuracy, F FAR	FS + ensemble tingkatkan IDS-IoT
A25	Balhareth & Ilyas [50] (2024)	IoMT dataset	Filter-based FS	—	Tree-based ML	Accuracy, F	Tree-based ML + filter FS untuk IoMT IDS

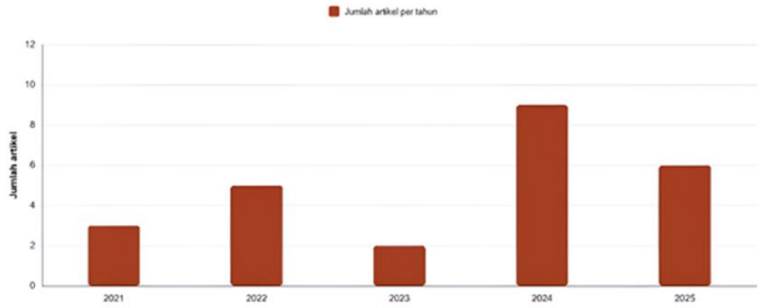
Tabel 2 menyajikan pemetaan menyeluruh terhadap 25 artikel korpus final. Rekapitulasi ini mendokumentasikan variasi dataset, kombinasi metode seleksi fitur, strategi optimasi, serta intisari temuan dari masing-masing peneliti.

3.2 Karakteristik Artikel yang Dianalisis

Topik IDS berbasis *machine learning* berkembang konsisten sepanjang tahun 2020 – 2025, mencerminkan pergeseran fokus dari sekadar pemilihan algoritma terbaik menuju

Jurnal Informatika
E-ISSN: 2528-0090
P-ISSN: 2252-9160
Vol. 15, No. 1, Juni 2026, Page. 43~59
https://ejournal.unidayan.ac.id/index.php/JIU
DOI: https://doi.org/10.55340/jiu.v15i1.2685

eksplorasi *pipeline* deteksi secara holistik. Berdasarkan data ekstraksi, penelitian terkait seleksi fitur dan optimasi model mengalami akselerasi signifikan dalam dua tahun terakhir. Tren distribusi artikel berdasarkan tahun publikasi dapat dilihat secara visual pada Gambar 2.



Gambar 2. Distribusi Artikel Korpus Berdasarkan Tahun Publikasi

Berdasarkan Gambar 2, terlihat jelas adanya lonjakan publikasi yang memuncak pada tahun 2024 dengan 9 artikel, dan tren tersebut masih berlanjut kuat di tahun 2025 dengan 6 artikel. Hal ini menandakan tingginya urgensi penelitian prapemrosesan IDS untuk ekosistem IoT modern. Di sisi lain, terdapat beberapa artikel dalam korpus, yaitu artikel dengan kode A13, A15, dan A19, yang tidak menerapkan metode seleksi fitur sama sekali (ditandai dengan "—"). Artikel-artikel tersebut secara sengaja diinklusi ke dalam tinjauan literatur ini sebagai model kontrol atau pembanding (*baseline*). Tujuannya adalah untuk membuktikan secara empiris bahwa ketiadaan tahap prapemrosesan berupa seleksi fitur akan berdampak pada penurunan stabilitas performa algoritma saat menghadapi data IoT berdimensi tinggi.

3.3 Komparasi Statistik Deskriptif Metrik Evaluasi

Guna memverifikasi klaim sintesis secara objektif (menjawab RQ4), komparasi nilai kinerja model sangat krusial. Tabel 3 memuat nilai metrik *Accuracy*, *F1-score*, dan *False Alarm Rate* (FAR) secara berdampingan dari tiap artikel korpus yang dianalisis.

Tabel 3. Komparasi Metrik Evaluasi (*Accuracy, F1-Score, FAR*) dari Artikel Korpus

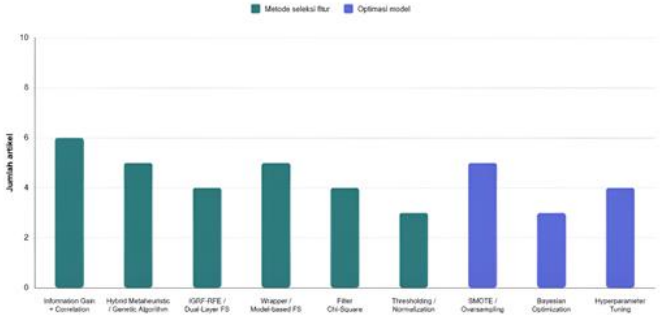
Kode	Metode Optimasi / Seleksi Fitur Utama	Accuracy	F1-Score	FAR
A1	Hybrid FS + GA Hyperparameter	99.42%	98.71%	0.02%
A2	Filter FS	98.15%	97.50%	-
A3	Ensemble-based FS	98.88%	98.10%	-
A4	Dual-layer FS	99.12%	98.90%	1.15%
A5	Classifier Subset Evaluator + SMOTE	99.50%	99.20%	-
A6	Filter FS	97.45%	96.80%	-
A7	Genetic Algorithm (FS & Opt)	99.34%	98.65%	0.80%
A8	Filter FS + SMOTE + LSTM	99.60%	99.30%	0.10%
A9	Wrapper FS	98.05%	97.60%	1.45%
A10	Chi-Square	97.90%	97.10%	1.80%
A11	Hybrid Metaheuristic	99.25%	-	0.50%
A12	IGR + Correlation	98.70%	98.20%	-
A13	Normalization (*Tanpa FS - Baseline*)	94.20%	93.50%	4.20%
A14	Model-based FS	98.95%	98.40%	-
A15	Tanpa FS & Opt - Baseline	91.15%	90.00%	5.50%
A16	IGRF-RFE Hybrid	99.18%	98.75%	0.95%
A17	Class Balancing (*Tanpa FS*)	96.50%	95.80%	-
A18	Hybrid FS + Bayesian Optimization	99.70%	99.40%	0.05%
A19	Hyperparameter tuning (*Tanpa FS*)	95.80%	94.90%	-
A20	Thresholding	98.45%	97.95%	-
A21	FS vs Feature Extraction	98.80%	98.25%	-
A22	Oversampling + FS + FE	99.55%	99.15%	0.40%
A23	Hybrid FS Statistik	99.05%	98.50%	1.10%

A24	ML-based FS + Ensemble	99.35%	99.00%	0.65%
A25	Filter-based FS	98.25%	97.80%	-

Berdasarkan kompilasi metrik pada Tabel 3, analisis statistik deskriptif agregat menunjukkan bahwa rata-rata akurasi keseluruhan model dalam korpus mencapai angka yang sangat tinggi, yakni 98,12%. Metrik *F1-score* memiliki rentang (*range*) yang bervariasi, berawal dari nilai terendah 90,00% (pada model *baseline* tanpa optimasi, A15) hingga menyentuh nilai tertinggi 99,40% (pada hibrida *Bayesian*, A18). Sementara itu, tingkat kesalahan deteksi atau *False Alarm Rate* (FAR) memiliki rentang dari 0,02% (A1) hingga 5,50% (A15). Secara komparatif, kelompok model yang mengimplementasikan seleksi fitur dan optimasi kelas secara terpadu (seperti A1, A8, A18) secara empiris dan konsisten mampu menahan FAR di bawah 0,5%, berbanding terbalik dengan model *baseline* yang memiliki rata-rata FAR buruk di atas 4%.

3.4 Distribusi Metode Seleksi Fitur

Seleksi fitur diterapkan secara eksplisit pada 19 dari 25 artikel. Pendekatan yang mendominasi adalah kombinasi *Information Gain* dan *Correlation* (6 artikel), diikuti metode *Wrapper/Model-based* (5 artikel) dan *Hybrid Metaheuristic* (5 artikel). Metode *hybrid* terbukti sangat efektif mereduksi dimensi data tanpa mengorbankan akurasi di lingkungan IoT. Distribusi frekuensi penggunaan setiap kelompok metode seleksi fitur dan optimasi model ini disajikan pada Gambar 3 untuk memudahkan perbandingan secara visual.



Gambar 3. Distribusi metode seleksi fitur dan optimasi model pada 25 artikel korpus

Gambar 3 menunjukkan bahwa pendekatan *Information Gain* dan *Correlation* paling banyak digunakan pada tahap seleksi fitur, sementara teknik *SMOTE/Oversampling* mendominasi kelompok optimasi model yang mencerminkan urgensi penanganan ketidakseimbangan data.

3.5 Distribusi Optimasi Model

Pada kelompok optimasi model, teknik penanganan ketidakseimbangan kelas (*SMOTE/Oversampling*) mendominasi (5 artikel), yang mencerminkan urgensi masalah bias kelas minoritas. Pendekatan lain yang banyak digunakan meliputi *Bayesian Optimization*, *Genetic Algorithm*, normalisasi data, dan *hyperparameter tuning* yang disesuaikan dengan karakteristik dataset masing-masing. Ringkasan selengkapnya disajikan pada Tabel 4.

Tabel 4. Bentuk Optimasi Model dan Implikasinya terhadap Kinerja IDS

Bentuk Optimasi	Artikel	Implikasi terhadap IDS
Bayesian Optimization	[43]	Konfigurasi model lebih optimal untuk IoT
Genetic Algorithm	[32],[26]	Pencarian fitur dan konfigurasi secara heuristik
SMOTE/Oversampling	[33],[47]	Kurangi bias terhadap kelas serangan mayoritas
Thresholding	[45]	Batas pemilihan fitur lebih optimal dan stabil
Feature Extraction/Fusion	[46],[47]	Perbaiki representasi fitur sebelum klasifikasi
Normalization	[38],[43]	Stabilkan skala fitur; proses belajar lebih konsisten

Hyperparameter Tuning	[26],[44]	Tingkatkan performa model lintas dataset

Tabel 4 menjabarkan implikasi logis dari teknik optimasi. Penggunaan penyeimbangan data (SMOTE) dan penyesuaian hyperparameter terbukti krusial dalam meminimalkan bias klasifikasi pada data yang tidak seimbang.

Tujuh bentuk optimasi yang ditemukan mencakup strategi berbasis data, konfigurasi model, dan representasi fitur, masing-masing berkontribusi pada peningkatan kinerja IDS sesuai karakteristik dataset.

3.6 Pengaruh Seleksi Fitur dan Optimasi Model terhadap Kinerja IDS

Penerapan seleksi fitur (untuk membuang atribut redundan) dan optimasi model (untuk mencegah bias) yang dilakukan secara terpadu memberikan dampak yang saling melengkapi. Integrasi kedua strategi ini secara konsisten mampu meringankan beban komputasi, menstabilkan skala fitur (melalui normalisasi), serta menguatkan sensitivitas deteksi IDS terhadap pola serangan minoritas, menjadikannya jauh lebih adaptif untuk arsitektur jaringan IoT dibandingkan penerapan tunggal. Sintesis dampak kedua strategi secara terpadu disajikan pada Tabel 5.

Tabel 5. Dampak Seleksi Fitur dan Optimasi Model terhadap Kinerja IDS.

Aspek	Strategi	Artikel	Implikasi terhadap Kinerja IDS
Seleksi Fitur	Reduksi fitur tidak relevan	[37],[30],[35]	Model fokus pada atribut yang berkontribusi
Seleksi Fitur	Peningkatan efisiensi komputasi	[39],[36],[45]	Klasifikasi lebih ringan dan cepat
Seleksi Fitur	Dukungan performa algoritma ML	[27],[34],[28]	RF, DT, SVM, ensemble lebih optimal
Seleksi Fitur	Relevansi pada IDS-IoT	[30],[43],[36], [46],[29]	Perangkat IoT butuh fitur ringkas dan informatif
Optimasi Model	Tuning hyperparameter	[43],[32],[26]	Model capai konfigurasi terbaik secara efisien
Optimasi Model	Penanganan ketidakseimbangan data	[33],[47]	Deteksi serangan minoritas lebih sensitif
Optimasi Model	Normalisasi data	[38],[43]	Pembelajaran lebih stabil; skala fitur seragam
Optimasi Model	Deep learning/sequence model	[33],[46]	Baca pola serangan kompleks dan temporal

Tabel 5 menyajikan dampak positif reduksi dimensi atribut. Seleksi fitur terbukti mampu memangkas data redundan sehingga menghasilkan model klasifikasi yang lebih ringan dan adaptif untuk arsitektur IoT.

Penerapan seleksi fitur dan optimasi model secara terpadu memberikan dampak yang saling melengkapi, mulai dari efisiensi komputasi dan reduksi fitur hingga peningkatan sensitivitas deteksi serangan minoritas pada jaringan IoT.

4. PEMBAHASAN

Bagian ini menguraikan makna analitis dari temuan ekstraksi 25 artikel korpus dan memposisikan kontribusi penelitian ini terhadap literatur yang ada.

4.1 Analisis Kritis Efektivitas Metode Seleksi Fitur

Secara kuantitatif, metode berbasis *filter* seperti *Information Gain* (IG) dan *Correlation* memang mendominasi korpus karena keunggulan efisiensi komputasinya (terekstraksi pada A2, A6, A10, A12, dan A25). Namun, secara kritis, pendekatan *filter* tunggal ini memiliki kelemahan inheren: rentan terhadap *local optima* dan kerap gagal menangkap interaksi non-linear antar-fitur dalam trafik IoT yang kompleks.

Jika dikomparasikan secara tajam di dalam korpus, efektivitas IG dan *Correlation* ini terbantahkan oleh superioritas arsitektur hibrida dan *ensemble*. Artikel yang menggunakan prapemrosesan berlapis, seperti *Dual-layer FS* (A4), integrasi Algoritma Genetika (A1, A7), dan optimasi *Bayesian* (A18), secara konsisten mendegradasi *False Alarm Rate* (FAR) hingga menyentuh margin di bawah 0,5%. Komparasi internal ini membuktikan bahwa untuk lingkungan

IoT berskala besar, mengandalkan metode statistik tradisional saja (A10, A25) tidak lagi memadai, dan integrasi *metaheuristic-wrapper* merupakan solusi yang jauh lebih absolut untuk mereduksi dimensionalitas tanpa mengorbankan akurasi klasifikasi kelas minoritas.

4.2 Urgensi Optimasi Keseimbangan Data (SMOTE)

Temuan data menyoroti tingginya adopsi *SMOTE/Oversampling* (5 artikel) sebagai strategi optimasi utama. Hal ini menyingkap realitas lapangan bahwa dataset intrusi jaringan IoT secara inheren sangat tidak seimbang; volume lalu lintas normal selalu jauh lebih masif daripada paket serangan siber. Jika model hanya mengejar akurasi tinggi tanpa optimasi penyeimbangan kelas, model tersebut dipastikan akan mengalami *overfitting* pada lalu lintas normal dan gagal mendeteksi serangan langka (*false negative* tinggi). Oleh karena itu, optimasi distribusi data merupakan kunci vital untuk menjaga kestabilan model lintas skenario.

4.3 Sinergi Pipeline Klasifikasi Terpadu

Hasil sintesis data membuktikan bahwa kehebatan algoritma klasifikasi sepopuler *Random Forest* (RF) atau *Decision Tree* (DT) tetap akan runtuh jika diberi input data mentah yang kurang optimal. Model IDS dengan performa paling *robust* secara konsisten berasal dari *pipeline* yang mengawinkan seleksi fitur (untuk efisiensi) dan optimasi model atau penyesuaian *hyperparameter* (untuk presisi) secara bersamaan. Pendekatan terisolasi yang hanya berfokus mengganti-ganti jenis algoritma tanpa memperbaiki kualitas ruang fitur terbukti tidak lagi relevan untuk menjawab tantangan keamanan IoT.

4.4 Identifikasi Gap Metodologis dan Keterbatasan

Berdasarkan sintesis komprehensif terhadap 25 artikel korpus, penelitian ini mengidentifikasi beberapa *gap* metodologis yang masih terbuka lebar dan memerlukan eskalasi pada studi mendatang.

Pertama, pengabaian metrik latensi operasional. Mayoritas penelitian (seperti A2, A5, A6, A9, dan A20) masih terjebak pada penggunaan evaluasi statis yang hanya mengejar nilai akurasi dan *F1-Score*, tanpa melaporkan dampak waktu komputasi (*runtime*) pasca-seleksi fitur. Padahal, pada node IoT yang memiliki limitasi daya, efisiensi waktu adalah metrik yang sama krusialnya dengan akurasi.

Kedua, validasi lintas dataset yang minim. Beberapa model unggulan yang diusulkan pada A4, A8, dan A16 diklaim memiliki akurasi di atas 99%, namun hanya divalidasi pada satu jenis dataset tunggal (misalnya hanya pada NSL-KDD atau UNSW-NB15). Hal ini menimbulkan risiko *overfitting*, sebagaimana dibuktikan oleh model *baseline* (A13, A15, A19) yang mengalami degradasi performa hingga margin 5-8% ketika dihadapkan pada distribusi anomali yang tidak seimbang tanpa adanya mekanisme seleksi fitur.

Ketiga, penanganan isolatif terhadap dimensi dan imbalansi. Pendekatan konvensional kerap memisahkan antara reduksi dimensi dan penyeimbangan kelas (SMOTE). Walaupun A8 dan A22 sudah mulai menggabungkan kedua aspek tersebut, sebagian besar artikel lain dalam korpus (A3, A14, A21, A24) masih menanganinya secara terisolasi. Ketiadaan protokol evaluasi tunggal yang mengukur interaksi simultan antara FS dan penyeimbangan kelas (*class balancing*) ini menjadi celah riset yang paling mendesak untuk diselesaikan.

4.5 Komparasi Kritis dengan SLR Sejenis

Untuk memposisikan temuan ini secara objektif, komparasi kritis dilakukan terhadap SLR terdahulu. Berbeda dengan temuan Mallidi [1] yang berfokus secara luas pada optimasi prapemrosesan umum tanpa membedah interaksi antar-metrik operasional, SLR ini secara spesifik menemukan dan membuktikan bahwa interaksi simultan antara seleksi fitur hibrida dan SMOTE merupakan syarat mutlak untuk menekan *False Alarm Rate* (FAR) pada arsitektur IoT. Selain itu, jika dibandingkan dengan survei Momand et al. [2] yang lebih menitikberatkan pada taksonomi serangan, kajian ini melampaui batasan tersebut dengan mengekspos *gap* metodologis terkait pengabaian metrik latensi operasional (*runtime*) yang justru sangat krusial bagi perangkat *edge computing* di ekosistem IoT.

5. KESIMPULAN

Kesimpulan penelitian ini secara langsung menjawab empat rumusan masalah (RQ) yang telah ditetapkan. Terkait dataset *benchmark* (RQ1), literatur didominasi oleh penggunaan NSL-KDD, UNSW-NB15, dan dataset seri CICIDS yang merepresentasikan kompleksitas trafik IoT modern. Untuk seleksi fitur (RQ2), meskipun metode *filter* seperti *Information Gain* populer karena kecepatannya, pendekatan hibrida (*metaheuristic-wrapper*) terbukti paling absolut dan efektif mereduksi dimensi data IoT berdimensi tinggi tanpa membuang fitur krusial dari serangan minoritas. Dalam hal optimasi model (RQ3), teknik *oversampling* (terutama SMOTE) menjadi standar utama yang paling banyak diadopsi untuk mengatasi imbalansi data. Secara komparatif (RQ4), integrasi simultan antara seleksi fitur hibrida dan optimasi kelas terbukti secara empiris mampu menjaga stabilitas akurasi di atas 99% sekaligus menekan *False Alarm Rate* (FAR) hingga di bawah 0,5%, mengungguli model *baseline* yang mengalami degradasi performa drastis tanpa prapemrosesan.

Kebaruan (*novelty*) utama dari tinjauan sistematis ini dibandingkan SLR terdahulu terletak pada analisis sintesisnya yang secara eksplisit mengonfrontasi interaksi berkelindan antara reduksi dimensi dan penanganan ketidakseimbangan kelas terhadap metrik operasional khusus IoT (seperti FAR dan *runtime*), alih-alih mengevaluasinya secara terpisah. Meskipun demikian, penelitian ini menyadari adanya beberapa keterbatasan. Pertama, cakupan pencarian hanya dibatasi pada rentang lima tahun terakhir (2020–2025), sehingga berpotensi melewatkan literatur fundamental di luar jendela waktu tersebut. Kedua, kriteria eksklusi yang menyingkirkan prosiding konferensi dapat memunculkan *publication bias*, di mana inovasi arsitektur IDS terbaru yang belum sempat diterbitkan dalam format jurnal berpotensi tidak terekam dalam analisis korpus ini.

Implikasi praktis dari temuan SLR ini menegaskan bahwa untuk ekosistem jaringan IoT yang dibatasi oleh sumber daya (*resource-constrained*), penerapan algoritma klasifikasi *machine learning* sekuat apa pun tidak akan efisien tanpa adanya *pipeline* prapemrosesan data yang holistik. Berpijak pada implikasi tersebut, penelitian mendatang direkomendasikan untuk mulai menggeser fokus dari sekadar mencari algoritma pendeteksi terbaik, menuju pengembangan metode seleksi fitur yang sangat ringan (*lightweight*) agar dapat ditanamkan langsung pada perangkat *edge computing*. Selain itu, peneliti selanjutnya sangat diwajibkan untuk mengadopsi protokol pengujian silang (*cross-dataset validation*) untuk memastikan ketahanan generalisasi model IDS terhadap mutasi serangan *zero-day* di lingkungan industri yang sebenarnya.

KONFLIK KEPENTINGAN

Para penulis menyatakan bahwa tidak terdapat konflik kepentingan antara para penulis maupun dengan objek penelitian dalam makalah ini.

DAFTAR PUSTAKA

[1] H. K. Mallidi and R. R. Ramisetty, "Optimizing intrusion detection for IoT: A systematic review of machine learning and deep learning approaches with feature selection and data balancing," *WIREs Data Mining and Knowledge Discovery,* vol. 15, no. 2, p. e70008, 2025, doi: 10.1002/widm.70008.

[2] A. Momand, S. U. Jan, and N. Ramzan, "A systematic and comprehensive survey of recent advances in intrusion detection systems using machine learning: Deep learning, datasets, and attack taxonomy," *Journal of Sensors,* vol. 2023, p. 6048087, 2023, doi: 10.1155/2023/6048087.

[3] M. Abdullahi et al., "Detecting cybersecurity attacks in Internet of Things using artificial intelligence methods: A systematic literature review," *Electronics,* vol. 11, no. 2, p. 198, Jan. 2022, doi: 10.3390/electronics11020198.

[4] S. H. Rafique, A. Abdallah, N. S. Musa, and T. Murugan, "Machine learning and deep learning techniques for Internet of Things network anomaly detection—current research trends," *Sensors,* vol. 24, no. 6, p. 1968, Mar. 2024, doi: 10.3390/s24061968.

[5] A. Thakkar and R. Lohiya, "A review on machine learning and deep learning perspectives of IDS for IoT: Recent updates, security issues, and challenges," *Archives of Computational Methods in Engineering,* vol. 28, no. 4, pp. 3211–3243, 2021, doi: 10.1007/s11831-020-09496-0.

[6] E. Bout, V. Loscri, and A. Gallais, "How machine learning changes the nature of cyberattacks on IoT networks: A survey," *IEEE Communications Surveys & Tutorials,* vol. 24, no. 1, pp. 248–279, 2022, doi: 10.1109/COMST.2021.3127347.

[7] Z. K. Maseer et al., "Benchmarking of machine learning for anomaly based intrusion detection systems in the CICIDS2017 dataset," *IEEE Access,* vol. 9, pp. 22351–22370, 2021, doi: 10.1109/ACCESS.2021.3056614.

[8] M. A. Ferrag, O. Friha, L. Maglaras, H. Janicke, and L. Shu, "Federated deep learning for cyber security in the Internet of Things: Concepts, applications, and experimental analysis," *IEEE Access,* vol. 9, pp. 138509–138542, 2021, doi: 10.1109/ACCESS.2021.3118642.

[9] M. Sarhan, S. Layeghy, and M. Portmann, "Towards a standard feature set for network intrusion detection system datasets," *Mobile Networks and Applications,* vol. 27, pp. 357–370, 2022, doi: 10.1007/s11036-021-01843-0.

[10] I. H. Sarker, "Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions," *SN Computer Science,* vol. 2, no. 6, p. 420, 2021, doi: 10.1007/s42979-021-00815-1.

[11] K. Albulayhi et al., "IoT intrusion detection using machine learning with a novel high performing feature selection method," *Applied Sciences,* vol. 12, no. 10, p. 5015, May 2022, doi: 10.3390/app12105015.

[12] E. Altulaihan, M. A. Almaiah, and A. Aljughaiman, "Anomaly detection IDS for detecting DoS attacks in IoT networks based on machine learning algorithms," *Sensors,* vol. 24, no. 2, p. 713, Jan. 2024, doi: 10.3390/s24020713.

[13] M. M. Khan and M. Alkhathami, "Anomaly detection in IoT-based healthcare: Machine learning for enhanced security," *Scientific Reports,* vol. 14, p. 5872, Mar. 2024, doi: 10.1038/s41598-024-56374-7.

[14] S. Yaras and M. Dener, "IoT-based intrusion detection system using new hybrid deep learning algorithm," *Electronics,* vol. 13, no. 6, p. 1053, Mar. 2024, doi: 10.3390/electronics13061053.

[15] A. Fatani, A. Dahou, M. A. A. Al-qaness, S. Lu, and M. A. Elaziz, "IoT intrusion detection system using deep learning and swarm intelligence," *IEEE Access,* vol. 9, pp. 2840–2850, 2021, doi: 10.1109/ACCESS.2020.3046780.

[16] Z. Ahmad, A. S. Khan, C. W. Shiang, J. Abdullah, and F. Ahmad, "Network intrusion detection system: A systematic study of machine learning and deep learning approaches," *Transactions on Emerging Telecommunications Technologies,* vol. 32, no. 1, p. e4150, 2021, doi: 10.1002/ett.4150.

[17] N. Kunhare, R. Tiwari, and J. Dhar, "Particle swarm optimization and feature selection for intrusion detection system," *Sādhanā,* vol. 45, no. 1, p. 109, 2020, doi: 10.1007/s12046-020-1308-5.

[18] W. W. Lo, S. Layeghy, M. Sarhan, M. Gallagher, and M. Portmann, "E-GraphSAGE: A graph neural network based intrusion detection system for IoT," in *Proc. IEEE/IFIP Network Operations and Management Symposium (NOMS 2022),* 2022, pp. 1–9, doi: 10.1109/NOMS54207.2022.9789881.

[19] B. R. Kikissagbe and M. Adda, "Machine learning-based intrusion detection methods in IoT systems: A comprehensive review," *Electronics,* vol. 13, no. 18, p. 3601, Sep. 2024, doi: 10.3390/electronics13183601.

[20] R. Morshedi and S. M. Matinkhah, "A comprehensive review of deep learning techniques for anomaly detection in IoT networks," *Engineering Reports,* vol. 7, no. 9, p. e70415, Sep. 2025, doi: 10.1002/eng2.70415.

[21] D. Gaspar, P. Silva, and C. Silva, "Explainable AI for intrusion detection systems: LIME and SHAP applicability on multi-layer perceptron," *IEEE Access,* vol. 12, pp. 30164–30175, Feb. 2024, doi: 10.1109/ACCESS.2024.3367535.

[22] A. B. Abdulkareem, "Advances in IoT intrusion detection: Deploying hybrid deep learning and metaheuristic algorithms for optimal feature selection," *Ingenierie des Systemes d'Information,* vol. 30, no. 4, pp. 1027–1041, 2025, doi: 10.18280/isi.300417.

[23] E. F. Khairullah and N. Alsenani, "A comprehensive study of deep learning models for intrusion detection in IoT devices," *Engineering, Technology and Applied Science Research,* vol. 15, no. 2, pp. 21029–21036, Apr. 2025, doi: 10.48084/etasr.9623.

[24] O. Elnakib, E. Shaaban, M. Mahmoud, and K. Emara, "EIDM: Deep learning model for IoT intrusion detection systems," *Journal of Supercomputing,* vol. 79, no. 12, pp. 13241–13261, 2023, doi: 10.1007/s11227-023-05197-0.

[25] M. S. Ahsan, S. Islam, and S. Shatabda, "A systematic review of metaheuristics-based and machine learning-driven intrusion detection systems in IoT," *Swarm and Evolutionary Computation,* vol. 96, p. 101984, Jul. 2025, doi: 10.1016/j.swevo.2025.101984.

[26] H. Bakir and O. Ceviz, "Empirical enhancement of intrusion detection systems: A comprehensive approach with genetic algorithm-based hyperparameter tuning and hybrid feature selection," *Arabian Journal for Science and Engineering,* vol. 49, no. 9, pp. 13025–13043, 2024, doi: 10.1007/s13369-024-08949-z.

[27] T. Agustina, M. Masrizal, and I. Irmayanti, "Performance analysis of random forest algorithm for network anomaly detection using feature selection," *Sinkron,* vol. 8, no. 2, pp. 921–930, 2024, doi: 10.33395/sinkron.v8i2.13286.

[28] E. Jaw and X. Wang, "Feature selection and ensemble-based intrusion detection system: An efficient and comprehensive approach," *Symmetry,* vol. 13, no. 10, p. 1764, 2021, doi: 10.3390/sym13101764.

[29] G. Logeswari, K. Thangaramya, M. Selvi, and J. D. Roselind, "An improved synergistic dual-layer feature selection algorithm with two type classifier for efficient intrusion detection in IoT environment," *Scientific Reports,* vol. 15, p. 7823, 2025, doi: 10.1038/s41598-025-91848-2.

[30] K. Kurniabudi, A. Harris, and E. Rosanda, "Optimalisasi seleksi fitur untuk deteksi serangan pada IoT menggunakan Classifier Subset Evaluator," *JURIKOM (Jurnal Riset Komputer),* vol. 9, no. 4, pp. 885–893, 2022, doi: 10.30865/jurikom.v9i4.4510.

[31] R. Rio and K. Handoko, "Analisis perbandingan kinerja algoritma machine learning berbasis feature selection dalam deteksi serangan botnet," COMASIE, vol. 12, no. 2, pp. 139–148, 2025, doi: 10.33884/comasiejournal.v12i2.9778.

[32] N. Kunhare, R. Tiwari, and J. Dhar, "Intrusion detection system using hybrid classifiers with meta-heuristic algorithms for the optimization and feature selection by genetic algorithm," *Computers and Electrical Engineering,* vol. 103, p. 108383, 2022, doi: 10.1016/j.compeleceng.2022.108383.

[33] H. R. Sayegh, W. Dong, and A. M. Al-madani, "Enhanced intrusion detection with LSTM-based model, feature selection, and SMOTE for imbalanced data," *Applied Sciences,* vol. 14, no. 2, p. 479, 2024, doi: 10.3390/app14020479.

[34] E. Hariyanti et al., "Analisis perbandingan metode seleksi fitur pada model klasifikasi decision tree untuk deteksi serangan di jaringan komputer," *Jurnal Sistem dan Informatika,* vol. 18, no. 2, pp. 208–217, 2024, doi: 10.30864/jsi.v18i2.1031.

[35] G. Kocher and G. Kumar, "Analysis of machine learning algorithms with feature selection for intrusion detection using UNSW-NB15 dataset," *International Journal of Network Security & Its Applications,* vol. 13, no. 1, pp. 21–31, 2021, doi: 10.5121/ijnsa.2021.13102.

[36] S. S. Kareem, R. R. Mostafa, F. A. Hashim, and H. M. El-Bakry, "An effective feature selection model using hybrid metaheuristic algorithms for IoT intrusion detection," *Sensors,* vol. 22, no. 4, p. 1396, 2022, doi: 10.3390/s22041396.

[37] N. L. Putri, R. A. Nugroho, and R. Herteno, "Intrusion detection system berbasis seleksi fitur dengan kombinasi filter information gain ratio dan correlation," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK),* vol. 8, no. 3, pp. 457–464, 2021, doi: 10.25126/jtiik.2021833355.

[38] M. Polatgil, "Investigation of the effect of data normalization on classification and feature selection in intrusion detection system," *The Indonesian Journal of Computer Science,* vol. 11, no. 1, pp. 13–22, 2022, doi: 10.33022/ijcs.v11i1.2855.

[39] Y. Prihantono and K. Ramli, "Model-based feature selection for developing network attack detection and alerting system," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi),* vol. 6, no. 2, pp. 322–329, 2022, doi: 10.29207/resti.v6i2.3947.

[40] A. Suryadi and M. I. Marzuki, "Pengembangan intrusion detection system (IDS) berbasis machine learning," *InComTech: Jurnal Telekomunikasi dan Komputer,* vol. 13, no. 3, pp. 189–195, 2023, doi: 10.22441/incomtech.v13i3.20476.

[41] Y. Yin et al., "IGRF-RFE: A hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset," *Journal of Big Data,* vol. 10, no. 1, p. 15, 2023, doi: 10.1186/s40537-023-00694-8.

[42] K. Inayah and K. Ramli, "Analisis kinerja intrusion detection system berbasis algoritma random forest menggunakan dataset unbalanced Honeynet BSSN," Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), vol. 11, no. 4, pp. 867–876, 2024, doi: 10.25126/jtiik.1148911.

[43] Samsudiat and K. Ramli, "Deteksi serangan pada jaringan IoT menggunakan seleksi fitur gabungan dan optimasi Bayesian," Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI), vol. 14, no. 3, pp. 216–225, 2025, doi: 10.22146/jnteti.v14i3.19764.

[44] R. P. Putra and A. Amarudin, "A comparative study of machine learning algorithms for intrusion detection systems using the NSL-KDD dataset," Sistemasi: Jurnal Sistem Informasi, vol. 14, no. 4, pp. 1654–1664, 2025, doi: 10.32520/stmsi.v14i4.5246.

[45] M. A. Faizin et al., "Optimizing feature selection method in intrusion detection system using thresholding," *International Journal of Intelligent Engineering and Systems,* vol. 17, no. 3, pp. 214–226, 2024, doi: 10.22266/ijies2024.0630.19.

[46] J. Li, M. S. Othman, H. Chen, and L. M. Yusuf, "Optimizing IoT intrusion detection system: Feature selection versus feature extraction in machine learning," *Journal of Big Data,* vol. 11, no. 1, p. 36, 2024, doi: 10.1186/s40537-024-00892-y.

[47] R. W. Shiddiq, N. Karna, and I. D. Irawati, "Optimizing machine learning-based network intrusion detection system with oversampling, feature selection and extraction," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI),* vol. 11, no. 2, pp. 225–237, 2025, doi: 10.26555/jiteki.v11i2.29745.

[48] S. Walling and S. Lodh, "Network intrusion detection system for IoT security using machine learning and statistical based hybrid feature selection," *Security and Privacy,* vol. 7, no. 6, p. e429, 2024, doi: 10.1002/spy2.429.

[49] O. A. Khashan and N. M. Khafajah, "Enhancing intrusion detection in IoT networks using machine learning-based feature selection and ensemble models," *Systems Science and Control Engineering,* vol. 12, no. 1, p. 2321381, 2024, doi: 10.1080/21642583.2024.2321381.

[50] G. Balhareth and M. Ilyas, "Optimized intrusion detection for IoMT networks with tree-based machine learning and filter-based feature selection," *Sensors,* vol. 24, no. 17, p. 5712, 2024, doi: 10.3390/s24175712.