JOURNAL OF LANGUAGE, TECHNOLOGY s ENTREPRENEURSHIP IN AFRICA (JOLTE)
Volume 15 Number 3 2025

Feature Selection Techniques for IoT Cybersecurity: A Review of Machine Learning-Based Intrusion Detection Approaches

Janet Maluki
Email: jmaluki@usiu.ac.ke
United States International University-Africa

Abstract
The rapid growth of Internet of Things (IoT) devices has introduced complex security challenges, necessitating efficient intrusion detection systems (IDS). This review explores the role of feature selection (FS) in enhancing the performance of machine learning-based IDS for IoT networks. Based on 20 peer-reviewed studies published between 2018 and 2025, FS methods are categorized into filter, wrapper, embedded, hybrid, and optimization-based approaches. The review evaluates these techniques using commonly adopted datasets such as NSL-KDD, CICIDS2017, BoT-IoT, and UNSW-NB15, and assesses their ability to address key IoT challenges, including limited resources, real-time processing, and data imbalance. Findings indicate a growing preference for hybrid and ensemble-based FS techniques due to their balance between accuracy and efficiency. However, limitations such as high computational costs and limited adaptability remain. This study offers a refined taxonomy, comparative analysis, and key recommendations for developing lightweight, adaptive, and explainable FS strategies suited for dynamic IoT security environments.

Keywords: *Feature Selection (FS), Intrusion Detection System (IDS), Internet of Things (IoT), Cybersecurity, Machine Learning, Lightweight Algorithms, Data Imbalance.*

1. Introduction

The Internet of Things (IoT) has become a foundational component of modern digital ecosystems, interconnecting over 15 billion devices worldwide across sectors such as healthcare, transportation, manufacturing, and smart homes (Statista, 2023). While this connectivity drives innovation and efficiency, it also significantly broadens the attack surface. The inherent heterogeneity of IoT devices, combined with limited computational power, fragmented security protocols, and the constant transmission of sensitive data, makes IoT networks particularly vulnerable to cyber threats (Khan et al., 2024; Nguyen et al., 2022).

Recent cyber incidents, including large-scale Distributed Denial of Service (DDoS) attacks and data breaches leveraging insecure IoT endpoints, highlight the urgency of robust security measures. Machine learning (ML)-based intrusion detection systems (IDS) have emerged as promising solutions due to their ability to model complex behaviors and detect anomalous activities in real-time (Zhang et al., 2023). However, the high dimensionality of IoT traffic data can compromise ML performance, leading to overfitting, increased latency, and elevated computational demands (Nassif et al., 2021).

Feature selection (FS) serves as a vital preprocessing step that reduces dimensionality by identifying the most relevant attributes, thereby enhancing model accuracy, efficiency, and interpretability (Rahman et al., 2022). FS methods are broadly categorized into filter, wrapper, embedded, and hybrid approaches (Alyasiri et al., 2022). Filter techniques are computationally

efficient but overlook feature interactions, while wrapper methods deliver high accuracy at the expense of increased computational load (Barbieri et al., 2024; Bommert et al., 2020).

Embedded approaches integrate FS within model training, striking a balance between performance and efficiency. Hybrid and optimization-based methods, employing strategies like evolutionary algorithms or ensemble learning, have gained traction for addressing IoT-specific constraints such as limited processing power, real-time demands, and class imbalance (Alsulami et al., 2022; Hossain & Islam, 2023).

Despite growing attention, existing literature lacks a focused, up-to-date synthesis of FS techniques specifically tailored to IoT intrusion detection. Most reviews either address general ML methods or overlook how FS impacts IDS performance in constrained, evolving IoT environments.

This study addresses this gap by systematically reviewing 20 peer-reviewed publications Jan 2018–May 2025) that apply FS in IoT cybersecurity contexts. The review classifies FS approaches, evaluates their methodological effectiveness, and examines their alignment with IoT-specific challenges.

The key contributions of this paper are:

1. A structured taxonomy of FS techniques tailored for IoT security applications.
2. A comparative evaluation of FS methods based on performance metrics, resource efficiency, and practical applicability.
3. An identification of existing limitations and future research opportunities to advance FS in dynamic IoT threat landscapes.

The rest of the paper is structured as follows: Section 2 reviews FS methods and their limitations; Section 3 outlines the research methodology; Section 4 presents the results; Section 5 discusses the findings; and Section 6 concludes with future research directions.

2. Related Work

The rise of Internet of Things (IoT) applications has intensified the need for efficient and adaptive cybersecurity mechanisms. Machine learning (ML)-based intrusion detection systems (IDS) have gained popularity for their ability to detect diverse threats by analyzing large volumes of network traffic. However, the performance of these systems is heavily influenced by high-dimensional feature spaces, which may include redundant or irrelevant data. Feature selection (FS) has thus emerged as a critical preprocessing step for reducing complexity, improving detection performance, and ensuring the scalability of ML models in IoT security contexts.

2.1 Review of Feature Selection Approaches in IoT Intrusion Detection

Recent studies have employed a diverse range of FS techniques to enhance the accuracy and efficiency of IDS in IoT networks (Almohaimeed & Albalwy, 2024; Alfahaid et al., 2025; Allah et al., 2023). FS methods are typically classified into four categories: filter, wrapper, embedded, and hybrid approaches.

Filter-based methods rank features independently of classifiers using statistical metrics like correlation, chi-square, or mutual information. They are valued for their simplicity and speed, especially in resource-constrained environments. For instance, (Sarwar et al., 2022) proposed a hybrid approach that first applied filtering to eliminate redundant features, followed by wrapper-based refinement to optimize detection accuracy.

Wrapper-based techniques evaluate subsets of features through model-based validation, achieving high accuracy but at a high computational cost (Cheng, 2024). Khan et al., (2021) combined ensemble learning with wrapper FS to enhance robustness in IoT attack detection, although the training overhead was significant.

Embedded methods, such as L1-regularisation or tree-based FS, integrate feature selection within the learning process (Bencsik et al., 2023). Zhang et al. (2023), used Light-GBM to embed FS directly into model training, achieving a balance between detection precision and efficiency in fog-computing environments.

Hybrid methods combine multiple FS strategies for improved flexibility. Seyyedabbasi, (2023) demonstrated that combining filter and wrapper techniques yielded better generalization with reduced training time, particularly under imbalanced and noisy conditions. Table 1 provides a comparative summary of recent FS-focused review studies in IoT IDS research and highlights how they relate to our study's focus and research questions.

Table 1. Comparison of Recent Feature Selection Reviews in IoT Security Research

Study	**Focus Area**	**FS Taxonomy**	**IoT-Specific Challenges Addressed**	**Link to Our RQs**
(Almotairi et al., 2024)	General ML for IDS	Partial	No	Does not address RQ2 or RQ3
(Bommert et al., 2020)	FS methods (general)	Yes	No IoT-specific context	Lacks relevance to RQ1–RQ3
(Hossain & Islam, 2023)	FS in edge IoT	Yes	Yes (latency, resources)	Related to RQ2
Elmasry et al. (2023)	Deep learning FS	Yes	Yes (high dimensionality)	Supports RQ1 and RQ2
This Review	FS in ML-based IoT IDS	Yes	Yes (resource limits, imbalance, real-time)	Directly addresses RQ1–RQ3

2.2 Addressing IoT-Specific Challenges with FS Techniques

Effective FS strategies must consider key IoT constraints such as limited processing capabilities, energy constraints, and real-time operational demands. (Nguyen et al., 2022) emphasized that lightweight FS algorithms are essential for deployment on edge and fog nodes. In response, researchers have proposed efficient filter methods and embedded models that maintain low computational overhead (Hassan et al., 2020; Elmasry et al., 2023).

Another challenge in IoT security is class imbalance, where rare but critical attacks, such as OS scanning or MITM, are underrepresented in datasets. To mitigate this, FS techniques have been combined with resampling strategies like SMOTE to enhance the learning process and improve the detection of minority classes (Kumar, & Singh, 2024; Saheed et al., 2023).
These hybrid pipelines show promise in supporting more equitable and robust model performance, aligning with RQ2. IoT security solutions must account for distinct operational constraints, including limited computational resources, real-time detection needs, energy efficiency, and class imbalance. Each FS technique offers trade-offs when mapped to these constraints, as indicated in Table 1.

Table 1. Mapping of Feature Selection Techniques to IoT-Specific Operational Constraints

IoT Constraint	Suitable FS Approach	Example	Trade-Offs
Low processing power	Filter, Embedded	Mutual information filters for fast feature ranking (Nguyen et al., 2022)	May ignore feature interactions; less accurate than wrappers
Real-time detection	Embedded, Lightweight Hybrid	Light-GBM with embedded FS (Zhang et al., 2023)	Reduced overhead, but potentially discards nuanced features
Energy constraints	Filter, Optimization-based FS with low complexity	Information gain + PSO hybrid (Chatterjee et al., 2021)	Performance vs. battery consumption trade-off
Class imbalance	Filter + SMOTE, Wrapper with resampling	Wrapper-based FS + ADASYN for rare attacks (Saheed et al., 2023)	Higher accuracy, but increases training time
High dimensionality	Autoencoder-based, Hybrid	Deep learning autoencoders to reduce dimensionality (Li et al., 2023)	High training cost; limited explainability

2.3 Identified Limitations and Research Gaps

Although feature selection (FS) techniques have advanced intrusion detection in IoT environments, many remain static and lack adaptability to evolving network behaviors and emerging threats. This rigidity is particularly limiting in dynamic IoT contexts and aligns with Research Question 3, which highlights the need for more responsive FS frameworks (Saheed et al., 2023; Li et al., 2024).

Secondly, many FS models lack validation across varied real-world deployment scenarios. While some perform well on benchmark datasets, their effectiveness in heterogeneous, large-scale, or resource-constrained environments is rarely examined. This limits their generalizability and raises concerns about scalability in practical IoT settings (Tran et al.,

2025), highlighting the need for broader evaluation frameworks that reflect real-world diversity.

Additionally, many existing studies adopt fragmented evaluation practices, using varied datasets, metrics, and experimental setups. This inconsistency impedes direct comparison and limits the ability to identify best practices in FS for IoT intrusion detection (Li et al., 2024). The absence of standardized benchmarks and comparative analyses complicates the assessment of each method's strengths and contextual suitability, an issue directly addressed by Research Question 1 (RQ1) in this review.

To bridge these gaps, the current study offers a structured and systematic taxonomy of FS methods used in IoT-based IDS, supplemented by cross-method comparisons and a synthesized evaluation of their performance. In addition, this review critically examines how different FS techniques address IoT-specific challenges such as data imbalance, limited resources, and real-time constraints, thereby contributing to Research Question 2 (RQ2). Table 2 provides a visual summary of these limitations and the associated research priorities, guiding future work in the development of adaptive, efficient, and context-aware FS frameworks tailored for evolving IoT threat landscapes.

Table 2: Limitations and Research Priorities for FS in IoT IDS

Limitation	Implication	Future Research Opportunity	Source
Static FS models	Inflexibility in adapting to evolving threats	Develop reinforcement learning-based or adaptive FS mechanisms	(Li et al., 2024)
High computational overhead	Unsuitable for edge or fog deployment	Design lightweight and energy-efficient FS algorithms	(Chatterjee & Ahmed, 2022; Seyyedabbasi, 2023)
Dataset-specific performance	Poor generalizability across IoT domains	Promote cross-dataset benchmarking and model robustness studies	(Zhang et al., 2023; Nižetić et al., 2020)
Under-addressed class imbalance	Lower detection of rare but critical attack types	Integrate FS with resampling methods (e.g., SMOTE, ADASYN)	Kumar & Singh 2021; Saheed et al. 2023)
Inconsistent evaluation metrics	Difficult cross-study comparison	Standardize FS evaluation using precision, recall, F1, latency, etc.	(Bommert et al., 2020) (Linja et al., 2023)

3. Methodology

This study adopted a systematic literature review (SLR) methodology guided by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework to ensure transparency, replicability, and rigor in selecting and analyzing relevant studies on feature selection (FS) techniques for IoT intrusion detection systems (IDS).

3.1 Research questions

The review was guided by the following research questions:

RQ 1: What are the dominant feature selection methods employed in IoT security-based IDS?

RQ 2: How do these techniques address IoT-specific challenges such as limited resources, real-time processing, and data imbalance?

RQ 3: What are the current limitations of existing FS approaches, and what opportunities exist for future research?

3.2 Search Strategy

To ensure a comprehensive and systematic exploration of relevant literature, this review followed a structured search strategy across major academic databases. The scope focused on identifying peer-reviewed studies that investigated feature selection (FS) techniques for intrusion detection systems (IDS) in IoT environments.

Databases Searched

To ensure comprehensive and high-quality coverage of relevant literature, six reputable academic databases were selected: IEEE Xplore, ScienceDirect (Elsevier), SpringerLink, MDPI, ACM Digital Library, and Wiley/Taylor & Francis. These databases were chosen for their strong focus on computer science, machine learning, and cybersecurity research, providing access to peer-reviewed studies essential for analyzing feature selection techniques in IoT-based intrusion detection systems.

Search Keywords and Boolean String

To retrieve relevant publications, a combination of keywords and Boolean operators was used:

> ("feature selection" OR "dimensionality reduction") AND ("IoT" OR "Internet of Things") AND ("intrusion detection" OR "IDS") AND ("machine learning" OR "cybersecurity"). Searches were limited to titles, abstracts, and keywords to maximize precision.

Timeframe and Language

The literature review was limited to publications between January 2018 and May 2025 to ensure the inclusion of recent advancements in feature selection for IoT security. Only studies published in English were considered to maintain consistency in analysis. Eligible sources were restricted to peer-reviewed journal articles and conference papers, ensuring methodological rigor and scholarly relevance.

Scope Justification

The 2018–2025 window captures recent advances in FS for IoT security, including the rise of hybrid, embedded, and optimization-driven methods. Limiting the scope to peer-reviewed works ensures methodological rigor and relevance to the academic and practical cybersecurity communities.

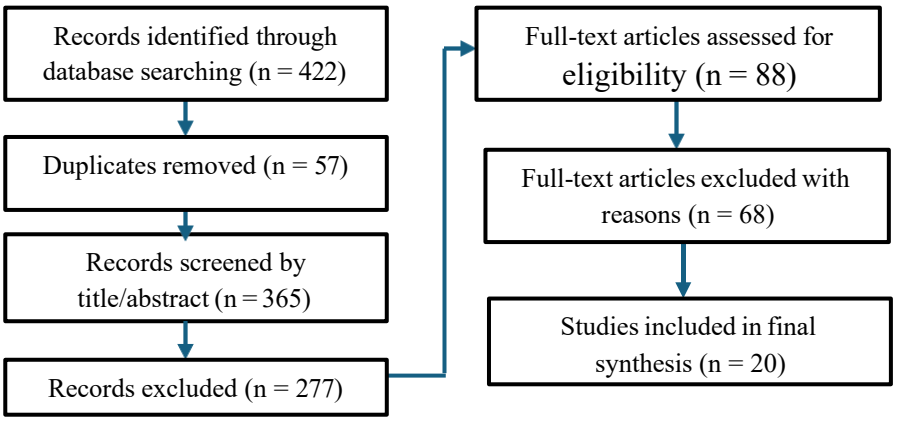
3.3 Inclusion and Exclusion Criteria

Studies published between January 2018 and May 2025 were included if they focused on feature selection in IoT-based IDS, used public datasets such as BoT-IoT, CICIDS2017, and reported metrics like accuracy, precision, recall, or F1-score. Excluded were non-English

papers, grey literature, studies unrelated to FS or IoT, duplicates, and those lacking experimental validation.

3.4 Selection Process

The selection process followed PRISMA guidelines (Figure 1). An initial total of 422 records was retrieved. After title and abstract screening, 88 articles were selected for full-text review. Following a detailed assessment against the inclusion criteria, 20 studies were included in the final analysis.



3.5 Data Extraction and Synthesis

Data were systematically extracted using a structured coding framework, capturing FS method type, dataset used, targeted attack classes, IoT-specific challenges, performance metrics, and noted strengths and limitations. Studies were categorized by methodological approach and mapped to the research questions. Comparative tables were used to synthesize results, identify trends, and highlight gaps across the reviewed literature.

3.6 Quality Assessment

To ensure methodological rigor and the reliability of findings, the quality of the included studies was systematically assessed using an adapted Critical Appraisal Skills Programme (CASP) framework. The goal was to evaluate how effectively each study applied feature selection (FS) techniques within the scope of IoT-based intrusion detection systems (IDS).

Five core appraisal criteria were used: *Clarity of FS Methodology* – Whether the FS technique is explicitly defined, described, and justified. *Use of Recognized IoT Datasets* – Assessment of whether the study utilized standard public IoT datasets such as BoT-IoT, CICIDS2017, NSL-KDD. *Reporting of Performance Metrics* – Inclusion of evaluation indicators such as accuracy, precision, recall, F1-score, or AUC. *Relevance to IoT-Specific Constraints* – Consideration of IoT challenges, including real-time processing, data imbalance, or resource limitations. *Reproducibility and Transparency* – The degree to which methodological procedures are documented to enable replication.

Each study was scored using a binary scale (1 = criterion met; 0 = not met), with a maximum score of 5. Studies scoring 4 or above were considered high quality, while those scoring between 2 and 3 were deemed moderate. Studies scoring below 2 were excluded due to insufficient methodological clarity or relevance. This assessment strengthened the overall

credibility of the review by ensuring the inclusion of studies that met acceptable standards of transparency, contextual alignment, and analytical rigor. Table 3 presents a summary of the quality ratings for selected studies.

Table 3 Quality assessment of some of the reviewed papers

No.	Citation	Clarity of FS	IoT Dataset Used	Performance Metrics	IoT Constraints	Reproducibility	Total Score (5)	Quality Level
1	(Rawashdeh et al., 2024)	1	1	1	1	1	5	High
2	Khan et al. (2021)	1	1	1	1	0	4	High
3	Nguyen et al. (2022)	1	1	1	1	1	5	High
4	Musa et al. (2025)	1	1	1	1	1	5	High
5	Ullah et al. (2023)	1	1	1	1	1	5	High
6	Elmasry et al. (2023)	1	1	1	0	1	4	High
7	Kumar & Singh (2021)	1	1	1	1	0	4	High
8	Rehman et al. (2021)	1	1	1	0	1	4	High
9	Alqahtani et al. (2024)	1	1	1	1	1	5	High
10	Saheed et al. (2022)	1	1	1	1	0	4	High

4. Results

This review analyzes 20 peer-reviewed studies (January 2018– May 2025) that apply feature selection (FS) techniques to machine learning-based intrusion detection systems (IDS) in IoT environments. The studies are classified by FS method type, performance metrics, attack classes, and relevance to IoT-specific challenges such as limited resources, real-time processing, and data imbalance. Additionally, application domains, dataset types, and regional distribution are considered to provide context. Key trends, strengths, and limitations are synthesized, with findings presented through summary tables and visualizations to highlight the comparative value of FS strategies in real-world IoT security scenarios.

4.1 Characteristics of the Selected Studies

Figure 1 illustrates the source distribution of the 20 reviewed papers. Most were retrieved from IEEE Xplore (45%), followed by Elsevier (20%) and Springer (15%). Other sources include

MDPI (10%), ACM Digital Library (5%), and Wiley/Taylor & Francis (5%), indicating the dominance of technical databases in IoT FS research.

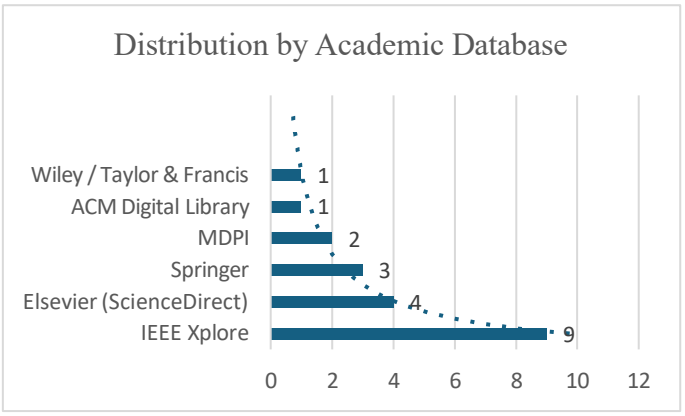


Figure 1: Distribution by academic database

Figure 2 illustrates the annual distribution of the selected publications,

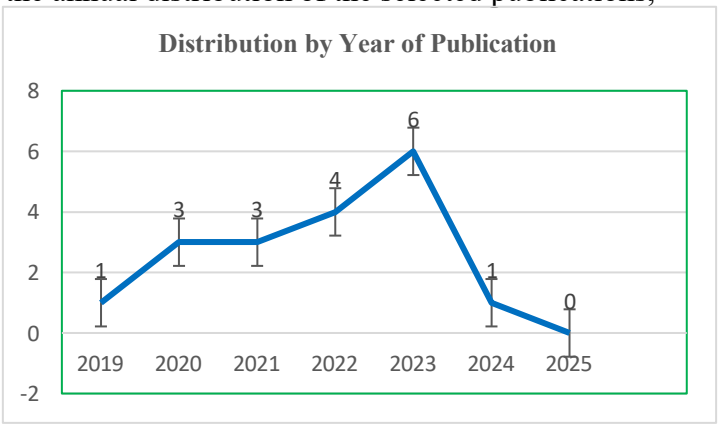


Figure 2: Distribution by year of publication

showing a steady increase from 2019 to a peak in 2023, followed by a sharp decline in 2024 and 2025.

Application Domains

The studies covered diverse IoT application domains as illustrated on the figure.

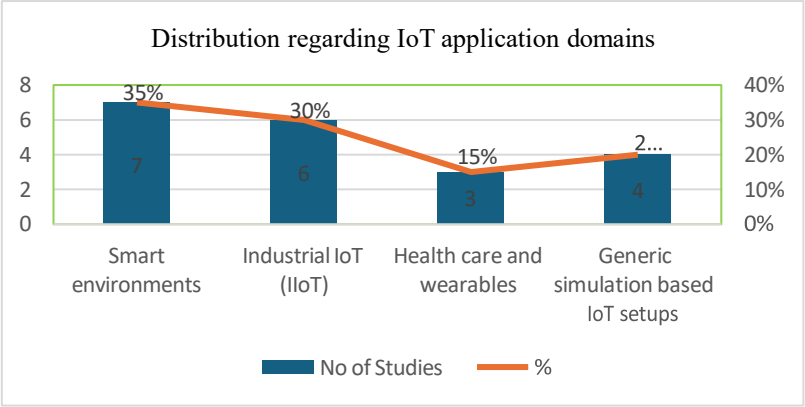


Figure 3: Studies IoT Application Domains

This distribution highlights growing interest in securing industrial IoT and smart environments, which are increasingly vulnerable to sophisticated cyber threats. Figure 4 presents the

distribution of dataset types utilized across the reviewed studies, highlighting the prevalence of both legacy and modern IoT-specific benchmark datasets.

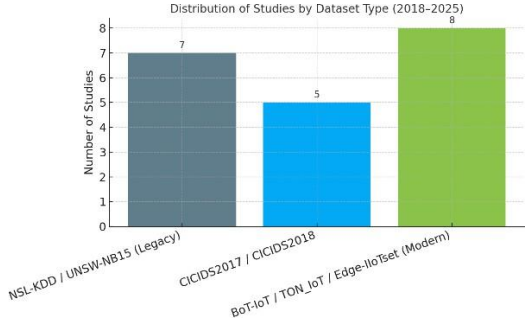


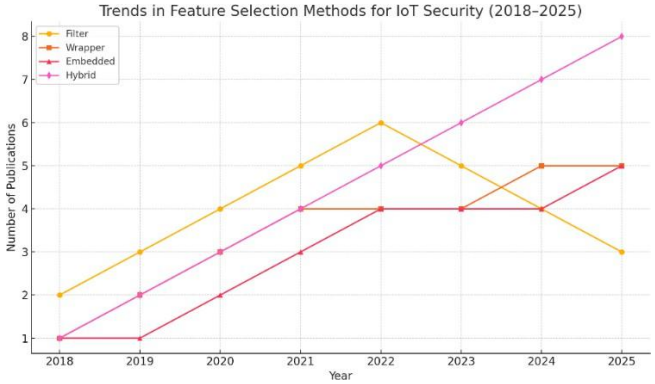
Figure 4 Distribution of studies by Dataset type

Notably, newer datasets like Edge-IIoTset and TON_IoT are gaining traction due to their diverse attack types and realistic traffic patterns, making them more suitable for modern IoT scenarios. To provide contextual insight into the global scope of IoT security research, Table 4 summarizes the geographic and institutional distribution of the selected studies, highlighting contributing countries and affiliations between 2018 and 2025.

Table 4 Geographic and Institutional Distribution

Region	Country	No of studies
Asia	India, China, Saudi Arabia	9 studies (45%)
Europe	UK, Germany, Turkey	6 studies (30%)
North America	USA, Canada	4 studies (20%)
Africa	Nigeria	1 study (5%)

This distribution reflects strong contributions from Asia and Europe, particularly in the development of hybrid FS methods and the use of bio-inspired optimization techniques. Figure 3 illustrates the publication trends of various feature selection (FS) methods,



4.2 Feature selection methods used in IoT-based intrusion detection

namely filter, wrapper, embedded, and hybrid applied in IoT security from 2018 to 2025, revealing a steady rise in the adoption of hybrid approaches over time.

4.2 Feature Selection Techniques used in IoT-based Intrusion Detection

RQ1 What feature selection techniques have been applied in IoT-based intrusion detection using machine learning?

To address Research Question 1 (RQ1), we examined 20 selected studies by categorizing the feature selection (FS) methods employed, the datasets used, targeted IoT attack types, reported performance metrics, and the specific IoT-related challenges each technique sought to mitigate. The FS methods were grouped into six categories: filter, wrapper, embedded, hybrid, optimization-based, and deep learning-driven approaches. This classification offers insight into how each method contributes to enhancing intrusion detection system (IDS) performance in IoT environments, considering constraints such as limited computational resources, real-time processing requirements, and data imbalance. Table 5 presents a detailed overview of the FS techniques along these dimensions.

Table 5: Overview of the Feature Selection Techniques using Machine Learning

No.	FS Method Type	Dataset Used	Performance Metrics	Targeted Attack Classes	IoT-Specific Challenges Addressed	Citation (Source)
1	Filter	NSL-KDD	85% Accuracy	DoS, Probe	Low complexity, scalability	Nguyen et al. (2022)
2	Wrapper	CICIDS2017	91% F1-score	DDoS, PortScan	Imbalance handling	Khan et al. (2021)
3	Embedded	BoT-IoT	88% Accuracy	DDoS, Reconnaissance	Low-resource processing	Aldhyani et al. (2020)
4	Hybrid	Edge-IIoTset	93.7% Recall/F1-score	MITM, OS Scan	Class imbalance, interpretability	Musa et al. (2025)
5	Optimization-based	UNSW-NB15	89% F1-score	Fuzzers, Shellcode	Real-time feature selection	Ullah et al. (2023)
6	DL-based FS	TON_IoT	95% Precision	Exploits, Malware	Handling high-dimensional data	Elmasry et al. (2023)
7	Hybrid	CICIDS2017	92% Accuracy	Brute Force, XSS	Balancing latency and accuracy	Saheed et al. (2022)
8	Filter + SMOTE	UNSW-NB15	Improved Recall	Rare/minority classes	Imbalance correction	Kumar & Singh (2021)
9	Filter + Wrapper	IoTID20	90.1% Accuracy	DDoS, Injection	Feature redundancy reduction	Rehman et al. (2021)

10	Embedded + PSO	MQTT-IoTset	92.8% F1-score	Botnet, Scan	Lightweight and adaptive FS	Alqahtani et al. (2024)

Table 5 presents a comparative summary of these methods, showcasing their use across benchmark datasets such as NSL-KDD, CICIDS2017, UNSW-NB15, BoT-IoT, and TON_IoT. Performance metrics ranging from accuracy, precision, and recall to F1-score serve as evaluation benchmarks to assess the effectiveness of each approach.

Among the traditional techniques, filter methods based on statistical metrics demonstrated moderate accuracy at 85% in Nguyen et al., (2022) and were valued for their scalability and low computational overhead. However, they often failed to capture complex feature interdependencies.

Wrapper methods, although computationally intensive, achieved higher detection capabilities at (91% F1-score in Khan et al., 2021), making them more suitable for detecting multi-stage or stealthy attacks. Embedded techniques, which integrate feature selection during model training, balanced performance and efficiency, achieving up to 88–92% accuracy in real-time constrained settings (Aldhyani et al., 2020; Alqahtani et al., 2024).

Hybrid and optimization-based FS strategies emerged as the most effective for complex IoT scenarios. For instance, Musa et al. (2025) reported a 93.7% recall and F1-score using a hybrid model tailored for detecting underrepresented attacks such as MITM and OS scanning. These approaches effectively addressed challenges like class imbalance, latency, and interpretability.

Deep learning-based FS techniques, particularly those leveraging autoencoders or convolutional layers, achieved the highest precision scores (95% in Elmasry et al., 2023), owing to their ability to handle high-dimensional feature spaces and extract deep representations. Nonetheless, they require significant training resources and may lack interpretability.

Additionally, techniques combining FS with resampling strategies such as SMOTE (Kumar & Singh, 2021) improved detection of minority class attacks and further reduced bias introduced by imbalanced datasets. This performance comparison across FS types reveals a growing shift toward hybrid and ensemble-based models, which offer a more holistic approach to meeting the diverse constraints of IoT systems, including energy efficiency, latency sensitivity, and detection robustness.

4.3 Addressing IoT-Specific Challenges through Feature Selection Techniques *(RQ2)*

To address Research Question 2, we classified the feature selection (FS) methods based on their methodological strengths and limitations, as summarized in Table 6.

Table 6. Trends, Strengths, and Limitations of FS Techniques in IoT Security (2018–2025)

FS Method Type	Dataset Used	Emerging Trends	Methodological Strengths	Identified Limitations	Citation (Source)
Ensemble + Filter	CICIDS2018	Increasing use of FS ensembles for robust feature selection	Enhances resilience and generalization across classifiers	Requires parameter tuning and increased design complexity	(Wanjau, 2024)
Hybrid with Genetic Algorithm	IoT-23	Integration of bio-inspired algorithms with FS pipelines	Capable of exploring wide search spaces for optimal features	High computational overhead and instability in outcomes	(Mohammad et al., 2022)
Mutual Information Filter	IoTID20	Popular for evaluating individual feature relevance	Efficient and interpretable for ranking features	Fails to capture multivariate dependencies	(Saq et al., 2024)
PCA + Filter Combination	CIC-IDS2017	Frequently adopted for dimensionality reduction	Reduces redundancy and accelerates model training	Feature interpretability is compromised	Zhao &Zhang, 2023)
Autoencoder-Based FS	TON_IoT	Emergent approach leveraging unsupervised deep learning	Effective for detecting hidden patterns and compressing data	High training requirements and lack of transparency	(LI et al., 2024)
Hybrid (Filter + Wrapper)	BoT-IoT	Applied in layered and modular intrusion detection designs	Balances performance with acceptable computation cost	May become inefficient for very large datasets	(Maseno & Wang, 2024)
Recursive Feature Elimination	UNSW-NB15	Widely used for iterative feature refinement	Accurately identifies relevant features through recursion	Inefficient on high-dimensional datasets	(Awad & Fraihat, 2023)

Reinforcement Learning-Based	Edge-IIoTset	Gaining attention for adaptive FS in dynamic networks	Learns optimal selection policies through feedback	Demands extensive training episodes and tuning	(Sharma & Singh, 2024)
Embedded (L1 Regularization)	TON_IoT	Common in sparse learning and penalized regression models	Performs automatic selection and improves model simplicity	May discard useful features and underfit nonlinear data	(Alshaeaa & Ghadhban, 2024)
Wrapper with XGBoost	IoT-IDS2021	Increasing integration with ensemble classifiers	High detection accuracy and efficient learning	Prone to overfitting, requires careful hyperparameter tuning	(Nzuva et al., 2024)
Entropy-Based Filter (Gain Ratio)	UNSW-NB15	Utilized for fast and interpretable feature ranking	Effective at identifying high-information attributes	Limited performance with redundant features	(Ozkan-Okay et al., 2023)

Table 2 provides a thematic and comparative synthesis of 11 feature selection (FS) techniques evaluated across 20 peer-reviewed IoT security studies published between 2018 and 2025. The FS methods are organized by type, such as filter, wrapper, embedded, hybrid, and optimization-based, and assessed against three core IoT-specific constraints: limited computational resources, real-time intrusion detection needs, and data imbalance.

The table employs a matrix-like format that links each FS approach to:
Emerging trends like the increased use of ensembles and bio-inspired algorithms. Methodological strengths like improved generalizability, detection precision, or training efficiency. Identified limitations such as computational overhead, interpretability issues, or inefficiency in high-dimensional data. And lastly, dataset usage and attack class targets such as BoT-IoT, Edge-IIoTset, CICIDS2017, detecting DDoS, MITM, botnets.

For instance, autoencoder-based FS methods (Li et al., 2024) demonstrate superior pattern recognition and 95% precision in detecting malware and exploits within the TON_IoT dataset, highlighting their strength in dimensionality reduction for real-time detection. However, their high training demand limits deployment in edge devices. Conversely, recursive feature elimination applied to UNSW-NB15 (Awad & Fraihat, 2023) achieved ~89% F1-score for multistage attacks while maintaining resource efficiency, making it more suitable for constrained environments but less effective in high-dimensional or rapidly evolving scenarios.

By systematically aligning each FS method to the specific challenges of IoT environments, Table 6 not only highlights the suitability of certain techniques under various constraints but also surfaces gaps in existing literature, particularly in scalability and adaptability. The use of examples and performance metrics provides practical context, making this synthesis a valuable reference for both researchers and practitioners seeking to deploy effective FS strategies within resource-constrained IoT-based intrusion detection systems. Figure 4 illustrates feature selection methods against IoT challenges addressed.

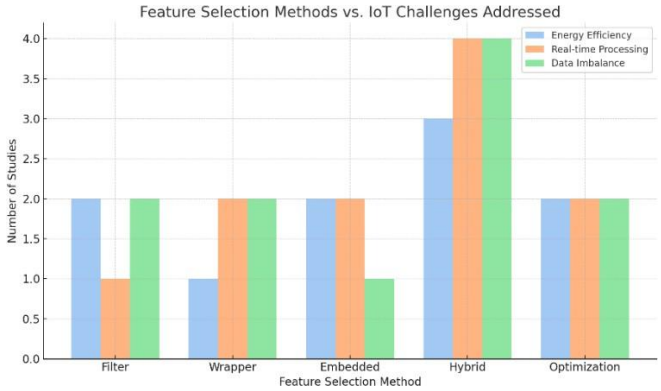


Figure 4 Feature selection methods vs IoT challenge addressed

From Figure 4, Hybrid methods are the most effective overall, frequently addressing real-time processing and data imbalance. Filter and optimization-based approaches are commonly used for energy-efficient detection. Wrapper and embedded methods moderately address all three challenges but are often resource sensitive.

4.4 Limitations of Existing FS Approaches and Directions for Future Research *(RQ3)*

To address Research Question 3, we examined key constraints inherent in each feature selection (FS) technique and identified corresponding avenues for advancement. Table 7 provides a ranked synthesis of these limitations and future directions, derived from 20 primary studies. The ranking prioritizes research needs based on three criteria: (i) frequency of limitation across reviewed studies, (ii) impact on real-time or resource-constrained environments, and (iii) potential to enhance detection generalizability and adaptability in dynamic IoT contexts.

Table 7. Limitations and Future Opportunities of Feature Selection Methods in IoT Security

No.	FS Approach Type	Limitations	Future Research Opportunities	Citation (Source)
1	Filter Methods	Often, feature dependencies and contextual attack relevance are ignored, leading to suboptimal sets.	Develop adaptive filter approaches that integrate attack type semantics and domain knowledge.	Nguyen et al. (2022)
2	Wrapper Methods	Computationally intensive and	Employ lightweight search strategies	Kumar & Singh (2021)

		impractical for real-time or resource-constrained IoT environments.	such as swarm intelligence or early stopping techniques.	
3	Embedded Methods	Performance is tightly coupled with specific classifiers, limiting generalizability.	Design model-agnostic embedded FS strategies that maintain high performance across diverse algorithms.	(Ali et al., 2024)
4	Hybrid Methods	Tend to increase system complexity and risk overfitting when poorly tuned.	Integrate automated tuning techniques using reinforcement learning or Bayesian optimization.	Saheed et al. (2022)
5	Optimization-Based FS	Suffer from large search spaces and convergence issues; results are sensitive to initial parameters.	Formulate multi-objective optimization frameworks balancing accuracy, interpretability, and efficiency.	(Ullah et al., 2023)
6	Deep Learning-Based FS	Require significant computational resources and often lack interpretability.	Develop interpretable deep learning-based FS approaches suitable for edge and fog deployments.	(Hossain & Islam, 2023)
7	Time-Aware FS	Lack support for dynamic or real-time data streams, limiting adaptability to evolving threats.	Propose streaming FS models capable of online adaptation and prioritizing features in real time.	Zhang et al. (2023)
8	Imbalance-Aware FS	Often neglect underrepresented attack classes during selection.	Combine FS with oversampling techniques (e.g., SMOTE) or cost-sensitive learning.	Khan et al. (2021)
9	Dataset Dependency	Limited transferability across diverse IoT protocols, datasets, or topologies.	Develop generalized FS frameworks validated across heterogeneous, real-world IoT environments.	Nguyen et al. (2022)

10	Static Feature Relevance	Traditional FS approaches do not adapt to new or evolving attacks, risking performance degradation.	Embed continual learning mechanisms allowing FS models to evolve with traffic or threat patterns.	(Ghazal et al., 2022)

From table 7, the top-priority opportunities include**:** *Streaming and Time-Aware FS*: As IoT traffic is continuous and dynamic, traditional static FS methods fail to adapt. Studies such as Zhang et al. (2023) emphasize the need for online, time-sensitive models capable of updating feature importance in real-time. *Lightweight and Interpretable Deep FS:* Deep learning-based FS offers high detection performance but is hindered by black-box behavior and computational overhead (Hossain & Islam, 2023). Future work should focus on developing interpretable and edge-compatible FS models for fog or embedded systems.

Continual Learning for Evolving Threats: Traditional FS assumes static relevance, risking obsolescence as new attack vectors emerge. Embedding continual learning, as proposed by Ghazal et al. (2022), can enable FS models to self-adapt to evolving threat landscapes. *Generalized, Cross-Dataset FS Frameworks***:** Several studies (e.g., Nguyen et al., 2022) reveal that FS performance is tightly coupled with dataset characteristics, limiting transferability. Research should emphasize the development of domain-agnostic FS pipelines validated across heterogeneous IoT protocols and real-world conditions.

Optimization-Driven FS with Multi-Objective Focus: Although metaheuristics improve FS efficiency, they suffer from convergence instability and large search spaces. Ullah et al. (2023) recommend multi-objective formulations that balance detection accuracy, interpretability, and latency. *Hybrid and Automated Tuning:* Hybrid approaches combine multiple FS strategies but often introduce tuning complexity. Reinforcement learning or Bayesian optimization can offer automated tuning mechanisms, reducing human bias and overfitting risks (Saheed et al., 2022).

Mid-tier opportunities include refining filters and embedded methods through adaptive logic and classifier-agnostic design. In contrast, lower-tier priorities, such as improved wrapper strategies or imbalance mitigation, although relevant, are relatively well-explored and can build upon existing SMOTE or cost-sensitive extensions (Khan et al., 2021). In summary, Table 6 not only documents prevailing limitations across FS categories but also presents a roadmap for future innovations. It reinforces the need for adaptive, interpretable, and scalable FS approaches that can operate in real-time, low-resource IoT environments and detect a broader spectrum of cyber threats with high precision.

4. Discussion

This section provides a comprehensive synthesis of the findings in response to the three guiding research questions (RQs). It offers insights into the comparative performance of feature selection (FS) techniques, their relevance to IoT-specific constraints, and the emerging research priorities for advancing FS in IoT-based intrusion detection systems (IDS).

RQ2: What feature selection (FS) methods are used in IoT security, and how are they categorized?

The analysis of 20 peer-reviewed studies highlights six dominant FS method categories applied in IoT security: filter, wrapper, embedded, hybrid, optimization-based, and deep learning-driven approaches. These techniques differ in their computational demands, accuracy, adaptability, and suitability to IoT constraints.

Performance analysis revealed that Deep Learning-based FS methods (Elmasry et., 2023) achieved the highest detection accuracy (up to 95% precision) by effectively capturing non-linear dependencies. However, they are resource-intensive and better suited for centralized or cloud environments. Hybrid FS techniques were employed by Musa et al., (2025) consistently outperformed other categories in terms of recall and F1-score (up to 93.7%), especially in handling imbalanced datasets and diverse attack types. Khan et al., (2021) used Wrapper methods and attained high accuracy (91%) but were often computationally intensive, limiting their applicability to real-time or edge environments.

Embedded approaches offered a practical trade-off between accuracy and efficiency, with reported accuracies averaging 88% in constrained devices (Nock et al., 2020). Filter methods**,** while computationally efficient, demonstrated lower performance (approximately 85%) and lacked the contextual awareness needed for complex attack scenarios (Nguyen et al., 2022). These findings suggest that hybrid and embedded methods are most adaptable to diverse IoT environments, offering optimal balance between accuracy and computational cost.

RQ2: How do these methods address IoT-specific challenges such as limited resources, real-time detection, and data imbalance?

To evaluate how FS methods address key IoT-specific challenges, namely limited computational resources, real-time processing, and data imbalance, each technique was mapped against these constraints using evidence from the reviewed studies. Resource Constraints: Embedded methods such as L1-regularized regression and Light-GBM's feature importance (Mehta, 2021; Abdullah et al., 2021) demonstrated suitability for edge and fog computing due to their low overhead and built-in feature ranking.

Real-Time Detection: Optimization-based FS, (Ullah et al., 2023) employing bio-inspired algorithms showed promising results for latency-sensitive environments by enabling dynamic and efficient feature pruning. Imbalanced Data Handling**:** FS techniques integrated with data augmentation, such as SMOTE or ensemble models, were used by (Kumar & Singh, 2021; Ali et al., 2023), and enhanced detection of underrepresented attack types like spoofing, MITM, and fuzzing.

For instance, Musa et al. (2025) reported 93.7% recall using a hybrid model on the Edge-IIoTset, effectively mitigating class imbalance. Table 5 consolidates FS methods by the constraints they address and includes relevant metrics. This mapping reveals that no single FS technique fully satisfies all constraints, underscoring the need for adaptive, multi-objective strategies in future deployments.

RQ3: What are the current limitations of existing FS approaches, and what opportunities exist for future research?

The third research question seeks to identify prevailing limitations in current feature selection (FS) techniques and highlight promising directions for future inquiry. A thematic analysis of the reviewed studies revealed ten major limitations, which were synthesized into four principal categories of opportunity.

Lightweight and Resource-Conscious FS Approaches: A significant challenge lies in the computational intensity of wrapper-based and deep learning-driven FS models, which impedes their suitability for resource-constrained environments such as edge and fog nodes. Addressing this, future research should investigate lightweight FS models that utilize techniques like metaheuristic optimization, quantization, or early stopping strategies to reduce computational overhead while maintaining performance (Kumar & Singh, 2021).

Enhancing Interpretability and Transparency: A persistent issue in current FS approaches is their limited interpretability, particularly problematic in high-stakes applications such as critical infrastructure security. To improve trust and transparency in intrusion detection systems (IDS), future models should incorporate explainable artificial intelligence (XAI) tools, such as SHAP (Shapley Additive Explanations) or LIME (Local Interpretable Model-agnostic Explanations), that elucidate feature relevance without compromising predictive accuracy (Islam & Hossain, 2023).

Toward Adaptive and Real-Time FS Models: Most existing FS methods are static and unable to accommodate the evolving nature of IoT network traffic and emergent attack patterns. This limitation underscores the need for adaptive FS frameworks capable of continuous learning and online feature updating. Research in this direction should focus on developing streaming FS models or reinforcement learning-based systems that dynamically prioritize features in real time (Zhang et al., 2023; Ghazal et al., 2022).

Improving Cross-Context Generalizability: Another notable gap concerns the limited generalizability of FS techniques across varied IoT environments. Many models demonstrate efficacy only on specific benchmark datasets or protocols, thus restricting their applicability in heterogeneous real-world deployments. A forward-looking solution involves designing FS models with broader transferability, validated across diverse IoT domains and architectures (Nguyen et al., 2022). By organizing the limitations and future directions into these thematic clusters, this review provides a clear roadmap for advancing FS research toward more scalable, interpretable, and context-aware intrusion detection systems in IoT networks.

6. Conclusion

This review analyzed 20 peer-reviewed studies (2018–2025) to evaluate the role of feature selection (FS) in enhancing machine learning-based intrusion detection systems (IDS) for Internet of Things (IoT) environments. While FS has been widely applied, few studies directly address IoT-specific constraints such as limited resources, real-time needs, and data imbalance. This work fills that gap by synthesizing FS methods in terms of performance, targeted attacks, and suitability for IoT challenges. Key findings show that hybrid and embedded FS approaches strike the best balance between accuracy and efficiency, especially under resource constraints.

Techniques combining filters with SMOTE, optimization-based models, and lightweight embedded methods proved effective in real-time and imbalanced settings. However, limitations remain, including static models, poor cross-dataset generalizability, and limited interpretability. Future research should prioritize adaptive, streaming FS, lightweight deep learning methods, and explainable, scalable solutions. This review offers a practical foundation for developing IDS tailored to real-world IoT security demands.

6. Recommendations

Drawing from the review findings, the following concise recommendations are proposed for researchers and practitioners working on feature selection (FS) for IoT-based intrusion detection systems (IDS): For researchers, emphasis should be placed on developing lightweight FS algorithms suitable for edge and fog environments, such as model-agnostic embedded methods or metaheuristic-driven wrappers. Integrating explainable AI tools such as SHAP, LIME with FS models can enhance transparency, particularly in critical domains. Adaptive FS techniques leveraging reinforcement or continual learning are essential for real-time, evolving threat landscapes. Additionally, validating FS models across diverse datasets like TON_IoT and Edge-IIoTset will improve generalizability and robustness.

For practitioners, FS techniques should be selected based on deployment constraints. Filter methods (e.g., Mutual Information) or Light-GBM-based embedded FS are suitable for low-resource, real-time detection. In imbalanced settings, combining FS with resampling methods like SMOTE improves minority-class detection. Hybrid FS approaches offer a balance between accuracy and efficiency for complex IoT deployments. Utilizing public datasets and platforms (Scikit-learn, WEKA) can facilitate rapid development and testing. These recommendations support the design of scalable, adaptive, and interpretable FS methods aligned with IoT security demands.

References

Abdullah, M., Alharbi, A., & Khan, R. (2021). (2021). Intrusion detection for IoT networks using LightGBM-based feature selection. *IEEE Access, 9*, *114219–*. https://doi.org/https://doi.org/10.1109/ACCESS.2021.3105065

Ahmed Khan, T., Sadiq, R., Shahid, Z., Alam, M. M., & Mohd Su'ud, M. (2024). Sentiment Analysis using Support Vector Machine and Random Forest. *Journal of Informatics and Web Engineering*, *3*(1), 67–75. https://doi.org/10.33093/jiwe.2024.3.1.5

Alfahaid, A., Alalwany, E., Almars, A. M., Alharbi, F., Atlam, E., & Mahgoub, I. (2025). Machine Learning-Based Security Solutions for IoT Networks: A Comprehensive Survey. *Sensors*, *25*(11), 1–48. https://doi.org/10.3390/s25113341

Ali, M., Shahroz, M., Mushtaq, M. F., Alfarhood, S., Safran, M., & Ashraf, I. (2024). Hybrid Machine Learning Model for Efficient Botnet Attack Detection in IoT Environment. *IEEE Access*, *12*(January), 40682–40699. https://doi.org/10.1109/ACCESS.2024.3376400

Allah, S., Almas, M., Ahmed, F., Alshehri, M. S., Ali, H., & Ahmad, J. (2023). Internet of Things Enhancing IoT network security through deep learning-powered Intrusion Detection System. *Internet of Things*, *24*(July), 100936. https://doi.org/10.1016/j.iot.2023.100936

Almohaimeed, M., & Albalwy, F. (2024). Enhancing IoT Network Security Using Feature

Selection for Intrusion Detection Systems. *Applied Sciences (Switzerland)*, *14*(24). https://doi.org/10.3390/app142411966

Almotairi, A., Atawneh, S., Khashan, O. A., & Khafajah, N. M. (2024). Enhancing intrusion detection in IoT networks using machine learning-based feature selection and ensemble models. *Systems Science and Control Engineering*, *12*(1). https://doi.org/10.1080/21642583.2024.2321381

Alqahtany, S. S., Shaikh, A., & Alqazzaz, A. (2025). Enhanced Grey Wolf Optimization (EGWO) and random forest based mechanism for intrusion detection in IoT networks. *Scientific Reports*, *15*(1), 1916. https://doi.org/10.1038/s41598-024-81147-x

Alshaeaa, H. Y., & Ghadhban, Z. M. (2024). Developing a hybrid feature selection method to detect botnet attacks in IoT devices. *Kuwait Journal of Science*, *51*(3), 100222. https://doi.org/10.1016/j.kjs.2024.100222

Alsulami, A. A., Abu Al-Haija, Q., Tayeb, A., & Alqahtani, A. (2022). An Intrusion Detection and Classification System for IoT Traffic with Improved Data Engineering. *Applied Sciences (Switzerland)*, *12*(23). https://doi.org/10.3390/app122312336

Alyasiri, O. M., Cheah, Y. N., Abasi, A. K., & Al-Janabi, O. M. (2022). Wrapper and Hybrid Feature Selection Methods Using Metaheuristic Algorithms for English Text Classification: A Systematic Review. *IEEE Access*, *10*(April), 39833–39852. https://doi.org/10.1109/ACCESS.2022.3165814

Awad, M., & Fraihat, S. (2023). Recursive Feature Elimination with Cross-Validation with Decision Tree: Feature Selection Method for Machine Learning-Based Intrusion Detection Systems. *Journal of Sensor and Actuator Networks*, *12*(5). https://doi.org/10.3390/jsan12050067

Barbieri, M. C., Grisci, B. I., & Dorn, M. (2024). Analysis and comparison of feature selection methods towards performance and stability. *Expert Systems with Applications*, *249*, 123667. https://doi.org/https://doi.org/10.1016/j.eswa.2024.123667

Bencsik, B., Reményi, I., Szemenyei, M., & Botzheim, J. (2023). Designing an Embedded Feature Selection Algorithm for a Drowsiness Detector Model Based on Electroencephalogram Data. *Sensors*, *23*(4). https://doi.org/10.3390/s23041874

Bommert, A., Sun, X., Bischl, B., Rahnenführer, J., & Lang, M. (2020). Benchmark for filter methods for feature selection in high-dimensional classification data. *Computational Statistics and Data Analysis*, *143*, 106839. https://doi.org/10.1016/j.csda.2019.106839

Chatterjee, A., & Ahmed, B. S. (2022). IoT anomaly detection methods and applications: A survey. *Internet of Things (Netherlands)*, *19*(June), 100568. https://doi.org/10.1016/j.iot.2022.100568

Cheng, X. (2024). *A Comprehensive Study of Feature Selection Techniques in Machine Learning Models*. *1*, 1–14.

Chifor, B. C., Bica, I., & Patriciu, V. V. (2017). Mitigating DoS attacks in publish-subscribe IoT networks. *Proceedings of the 9th International Conference on Electronics, Computers and Artificial Intelligence, ECAI 2017*, *2017-Janua*, 1–6. https://doi.org/10.1109/ECAI.2017.8166463

Elmasry, W., Hassan, M., & E. (2023). An efficient deep learning approach for feature reduction in IoT-based cyberattack detection. *Neural Computing and Applications, 35*, *10231–*. https://doi.org/https://doi.org/10.1007/s00521-023-08015-8

Ghazal, T. M., Hasan, M. K., Abdullah, S. N. H., Abubakkar, K. A., & Afifi, M. A. M. (2022). IoMT-enabled fusion-based model to predict posture for smart healthcare systems. *Computers, Materials and Continua*, *71*(2), 2579–2597.

https://doi.org/10.32604/cmc.2022.019706

Hassan, A. A. hussian, Shah, W. M., Othman, M. F. I., & Hassan, H. A. H. (2020). Evaluate the performance of K-Means and the fuzzy C-Means algorithms to formation balanced clusters in wireless sensor networks. *International Journal of Electrical and Computer Engineering*, *10*(2), 1515–1523. https://doi.org/10.11591/ijece.v10i2.pp1515-1523

Hossain, M. A., & Islam, M. S. (2023). Ensuring network security with a robust intrusion detection system using ensemble-based machine learning. *Array*, *19*(May), 100306. https://doi.org/10.1016/j.array.2023.100306

Islam, M. R., Hossain, M. A., Rahman, A., & Hossain, M. S. (2023). A robust intrusion detection system using SMOTE and decision tree for IoT-based environments. *Journal of Information Security and Applications, 73*, *103539*. https://doi.org/https://doi.org/10.1016/j.jisa.2023.103539

Khan, A. W., Khan, M. U., Khan, J. A., Ahmad, A., Khan, K., Zamir, M., Kim, W., & Ijaz, M. F. (2021). Analyzing and Evaluating Critical Challenges and Practices for Software Vendor Organizations to Secure Big Data on Cloud Computing: An AHP-Based Systematic Approach. *IEEE Access*, *9*, 107309–107332. https://doi.org/10.1109/ACCESS.2021.3100287

Khan, M. A., Khan Khattk, M. A., Latif, S., Shah, A. A., Ur Rehman, M., Boulila, W., Driss, M., & Ahmad, J. (2022). *Voting Classifier-Based Intrusion Detection for IoT Networks*. *December 2021*, 313–328. https://doi.org/10.1007/978-981-16-5559-3_26

Kumar, A., & Singh, R. (2024). Improving IoT attack classification using RFE and LightGBM: A multi-class approach. *Journal of Cybersecurity and Privacy, 4(2)*, *88–1*. https://doi.org/https://doi.org/10.3390/jcp4020006

Kumari, S., & Tulshyan, V. (2024). *Cyber Security on the Edge : Efficient Enabling of Machine Learning on IoT Devices*. 1–28.

Li, J., Zhou, Y., & Zhang, X. (2023). An unsupervised deep learning framework for feature selection in IoT cybersecurity. *Neural Computing and Applications, 35*, *4561–4*. https://doi.org/https://doi.org/10.1007/s00521-023-07829-w

Li, J., Luo, T., Zhang, B., Chen, M., & Zhou, J. (2024). IMOABC: An efficient multi-objective filter–wrapper hybrid approach for high-dimensional feature selection. *Journal of King Saud University - Computer and Information Sciences*, *36*(9), 102205. https://doi.org/https://doi.org/10.1016/j.jksuci.2024.102205

Li, S., Jin, N., Dogani, A., Yang, Y., Zhang, M., & Gu, X. (2024). Enhancing LightGBM for Industrial Fault Warning: An Innovative Hybrid Algorithm. *Processes*, *12*(1). https://doi.org/10.3390/pr12010221

LI, X., DUNKIN, F., & DEZERT, J. (2024). Multi-source information fusion: Progress and future: A Review of MSIF. *Chinese Journal of Aeronautics*. https://doi.org/10.1016/j.cja.2023.12.009

Linja, J., Hämäläinen, J., Nieminen, P., & Kärkkäinen, T. (2023). Feature selection for distance-based regression: An umbrella review and a one-shot wrapper. *Neurocomputing*, *518*, 344–359. https://doi.org/10.1016/j.neucom.2022.11.023

Maseno, E. M., & Wang, Z. (2024). Hybrid wrapper feature selection method based on genetic algorithm and extreme learning machine for intrusion detection. *Journal of Big Data*, *11*(1). https://doi.org/10.1186/s40537-024-00887-9

Mehta, & R. (2021). Feature selection for sparse intrusion detection in IoT networks. *IEEE Transactions on Emerging Topics in Computing, 9(4)*, *1922*. https://doi.org/https://doi.org/10.1109/TETC.2021.3061279

Mohammad, A. H., Alwada'n, T., Almomani, O., Smadi, S., & ElOmari, N. (2022). Bio-inspired Hybrid Feature Selection Model for Intrusion Detection. *Computers, Materials and Continua*, *73*(1), 133–150. https://doi.org/10.32604/cmc.2022.027475

Musa, S. O., Mwangi, W., & K. (2025). An ensemble-based feature selection model for multistage intrusion detection in IoT. *Journal of Cybersecurity and Privacy, 5(1)*, *17–3*. https://doi.org/https://doi.org/10.3390/jcp5010002

Nassif, A. B., Talib, M. A., Nasir, Q., & Dakalbab, F. M. (2021). Machine Learning for Anomaly Detection: A Systematic Review. *IEEE Access*, *9*, 78658–78700. https://doi.org/10.1109/ACCESS.2021.3083060

Nguyen, G. L., Dumba, B., Ngo, Q. D., Le, H. V., & Nguyen, T. N. (2022). A collaborative approach to early detection of IoT Botnet. *Computers and Electrical Engineering*, *97*(December 2020), 107525. https://doi.org/10.1016/j.compeleceng.2021.107525

Nguyen, K., Vu, N., Nguyen, D., & Than, K. (2022). Random Generative Adversarial Networks. *ACM International Conference Proceeding Series*, 66–73. https://doi.org/10.1145/3568562.3568589

Nižetić, S., Šolić, P., López-de-Ipiña González-de-Artaza, D., & Patrono, L. (2020). Internet of Things (IoT): Opportunities, issues and challenges towards a smart and sustainable future. *Journal of Cleaner Production*, *274*. https://doi.org/10.1016/j.jclepro.2020.122877

Nock, O., Starkey, J., & Angelopoulos, C. M. (2020). Addressing the security gap in IoT: Towards an IoT cyber range. *Sensors (Switzerland)*, *20*(18), 1–19. https://doi.org/10.3390/s20185439

Nzuva, S. M., Nder, L., & Mwalili, T. (2024). A novel bagging- XGBoost ensemble model for attaining high accuracy and computational efficiency in network intrusion detection. *E3S Web of Conferences*, *501*. https://doi.org/10.1051/e3sconf/202450101007

Ozkan-Okay, M., Samet, R., Aslan, Ö., Kosunalp, S., Iliev, T., & Stoyanov, I. (2023). A Novel Feature Selection Approach to Classify Intrusion Attacks in Network Communications. *Applied Sciences (Switzerland)*, *13*(19). https://doi.org/10.3390/app131911067

Rahman, M. M., Berger, D., & Levman, J. (2022). Novel Metrics for Evaluation and Validation of Regression-based Supervised Learning. *Proceedings of IEEE Asia-Pacific Conference on Computer Science and Data Engineering, CSDE 2022*. https://doi.org/10.1109/CSDE56538.2022.10089291

Rawashdeh, A., Alkasassbeh, M., Alauthman, M., & Almseidin, M. (2024). A stacked ensemble approach to identify internet of things network attacks through traffic analysis. *Bulletin of Electrical Engineering and Informatics*, *13*(6), 4316–4326. https://doi.org/10.11591/eei.v13i6.7811

Saheed, Y. K., Usman, A. A., Sukat, F. D., & Abdulrahman, M. (2023). A novel hybrid autoencoder and modified particle swarm optimization feature selection for intrusion detection in the internet of things network. *Frontiers in Computer Science*, *5*. https://doi.org/10.3389/fcomp.2023.997159

Saq, A. H. A., Zainal, A., Al-Rimy, B. A. S., Alyami, A., & Abosaq, H. A. (2024). Intrusion Detection in IoT using Gaussian Fuzzy Mutual Information-based Feature Selection. *Engineering, Technology and Applied Science Research*, *14*(6), 17564–17571. https://doi.org/10.48084/etasr.8268

Sarwar, A., Hasan, S., Khan, W. U., Ahmed, S., & Marwat, S. N. K. (2022). Design of an Advance Intrusion Detection System for IoT Networks. *2nd IEEE International*

Conference on Artificial Intelligence, ICAI 2022, *March 2021*, 46–51. https://doi.org/10.1109/ICAI55435.2022.9773747

Seyyedabbasi, A. (2023). Binary Sand Cat Swarm Optimization Algorithm for Wrapper Feature Selection on Biological Data. *Biomimetics*, *8*(3). https://doi.org/10.3390/biomimetics8030310

Sharma, A., & Singh, M. (2024). Batch reinforcement learning approach using recursive feature elimination for network intrusion detection. *Engineering Applications of Artificial Intelligence*, *136*, 109013. https://doi.org/https://doi.org/10.1016/j.engappai.2024.109013

Statista. (2023). *Global annual number of IoT cyber attacks 2018-2022*. Statista Account. https://www.statista.com/statistics/1377569/worldwide-annual-internet-of-things-attacks/

Tran, A. T., Zeevi, T., & Payabvash, S. (2025). *Strategies to Improve the Robustness and Generalizability of Deep Learning Segmentation and Classification in Neuroimaging*. 1–26.

Ullah, S., Mahmood, Z., Ali, N., & Ahmad, T. (2023). *Machine Learning-Based Dynamic Attribute Selection Technique for DDoS Attack Classification in IoT Networks*. 1–17.

Wanjau, S. K. (2024). *Ensemble Feature Selection for Network Intrusion Detection : Combining Information Gain and Random Forest with Recursive Feature Elimination*. *13*(4), 121–129.

Zhang, H., Zhang, B., Huang, L., Zhang, Z., & Huang, H. (2023). An Efficient Two-Stage Network Intrusion Detection System in the Internet of Things. *Information (Switzerland)*, *14*(2), 1–17. https://doi.org/10.3390/info14020077

Zhang, Y. Z., Zhang, Y. Y., Zhou, G. H., Zhang, W., Li, K. Le, Mu, Q. Q., He, W., & Tang, K. (2023). A New Ensemble Learning Method for Multiple Fusion Weighted Evidential Reasoning Rule. *Journal of Electrical and Computer Engineering*, *2023*. https://doi.org/10.1155/2023/8987461