

Use of Feature Extraction and Importance Techniques in Network Intrusion Detection Systems

OPEN ACCESS

Received: 20/12/2025

Accepted: 14/06/2026

Published: 07/07/2026

Sujata N Bhosle^{1*}, Jayshri D Pagare¹

¹ Jawaharlal Nehru Engineering College, MGM University, Chh.Sambhajinagar, Maharashtra, India

Abstract

Citation: Bhosle SN, Pagare JD (2026) Use of Feature Extraction and Importance Techniques in Network Intrusion Detection Systems. Indian Journal of Science and Technology 19(23): 1583-1592. <https://doi.org/10.17485/IJST/v19i23.1966>

* **Corresponding author.**

magaresujata@rediffmail.com

Funding: None

Competing Interests: None

Copyright: © 2026 Bhosle & Pagare. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](#))

ISSN

Print: 0974-6846

Electronic: 0974-5645

Background/Objectives: Feature Extraction plays an important role in every classification mechanism. The primary objective of this research is to systematically analyze and organize recent advancements from the year 2020-2025 in feature extraction, selection, and explainable AI techniques used for network traffic classification, and to find out the research gap that exists in existing work. Also, to evaluate existing feature extraction approaches available in NIDS and examine their limitations. **Method:** This review examines the main techniques used to determine feature importance in the context of network attack detection, ranging from simple permutation tests to model-independent explanation tools and newer game-theory-based approaches such as SHAP. Their ability to capture interactions among features, support real-time analysis, and limit issues like overfitting and bias are discussed in detail. **Findings:** Overall, the review shows that Feature Extraction methods play a practical role in building more transparent, reliable, and effective network traffic classification systems for detecting malicious activity. Feature-importance methods (SHAP, LIME, filter scores, tree importance) are widely applied to improve and explain network-attack classifiers, with studies showing large performance gains from targeted feature selection and XAI-enabled trust. Practical gaps remain in standardization, dataset biases, and packet-vs-flow attribution. **Novelty:** This study presents a comprehensive synthesis explicitly integrating feature extraction, feature selection, feature importance, Explainable Artificial Intelligence (XAI), and feature correlation analysis for Network Intrusion Detection Systems. Through a systematic review of recent studies (2020–2025), with particular emphasis on modern datasets such as NF-ToN-IoT-v2, this review identifies critical research gaps, proposes a correlation-aware explainable feature-engineering framework, and outlines a roadmap for developing transparent, scalable, and trustworthy intrusion detection systems. Our comparative study will help the researcher to choose the right feature extraction technique for their study.

Keywords: Feature-importance; Correlation Analysis; SHAP; LIME; Filter scores; Tree importance; NIDS

1 Introduction

With the rapid development of smart city-oriented network devices diverse data are produced, which open up new possibilities for modern attacks. Through effective network traffic classification, we can eliminate the possibility of a network attacks⁽¹⁾. Identifying attacks within network traffic increasingly relies on machine learning, yet these models are often suffers to understand which feature contribute most. Feature importance methods help overpass this gap by showing which traffic characteristics control a model's decisions and by helping reduce irrelevant features⁽²⁾.

Traditional port-based methods are not inadequate to work, due to today's dynamic nature of port. Machine learning and deep learning algorithm appeared as a prevailing research paradigm due to their capability to learn complex sample from high-dimensional netflow data and generalize to real-world attacks⁽³⁾.

The Performance of a machine learning algorithm depends upon the quality of selected features⁽⁴⁾. This will have an impact on the detection result, like accuracy and scalability of the model. Network traffic data is highly dimensional, correlated and diverse in nature⁽⁵⁾. Studies from 2020 are focusing on feature importance, feature extraction, and feature selection techniques to improve the performance of the model and to reduce the computational overhead^{(6), (7)}.

Statistical and protocol-level feature extraction techniques showed the impact on early netflow based NIDS datasets. Filter-based techniques such as Information Gain, Correlation, Mutual Information, and Chi-square were typically used on popular NIDS datasets i.e. CIC-IDS2017, CSE-CIC-IDS2018, and UNSW-NB15, to improve model's performance⁽⁸⁾.

With the escalating deep learning, studies progressively adopted automated feature extraction methods, including long short-term memory (LSTM), convolution neural networks (CNNs), recurrent neural networks (RNNs), gated recurrent units (GRU), and attention-based architectures^{(9), (10)}. These models learn hierarchical and spatiotemporal feature representations directly from raw or modestly processed flow data, achieving better performance to tackle recent attacks such as ransomware, DDoS, and zero-day exploits⁽¹¹⁾. For Dimensionality reduction in IoT environments, Auto-encoder-based mechanisms have been shown to be effective⁽¹²⁾.

Recent studies have proposed a hybrid approach to provide stable accuracy and low computational cost by combining feature extraction with feature selection methods along with ensemble learning^{(13), (14), (15)}. However, while these approaches improve predictive results, they often function as opaque models, restraining their implementation in operational and risk sensitive environment.

Feature importance in network security has progressed from simple filter scores and tree importance to model-agnostic XAI methods that provide local and global attributions for complex models. This evolution aims to both boost detection performance and make opaque models verifiable for operational use. Feature selection and explanation are shown to significantly improve IDS detection accuracy and reduce feature dimensionality, enabling near-peak accuracy with far fewer features in controlled environments⁽¹⁶⁾. Explainability has been integrated to increase conviction without reducing detection performance in ensemble classifiers and intrusion detection models on benchmark datasets^{(17), (18), (19), (20), (21)}.

1.1 Key Challenges In Network Traffic Classification

As the data grows in the smart city, it has to tackle multiple challenges with respect to Data, Model and and Operational challenges.

1.1.1. Data Challenges

- Class Imbalance: Normal traffic vastly outnumbers attack traffic (ratios of 100:1 or higher).
- Concept Drift: Attack patterns evolve, requiring adaptive feature importance.
- High Dimensionality: Modern network features can exceed 100+ dimensions.
- Temporal Dependencies: Packet sequences and flow dynamics require specialized handling.

1.1.2. Model Challenges

- Accuracy vs. Interpretability Trade-off: Deep models achieve highest accuracy but are least interpretable.
- Computational Cost: Real-time explanation generation can introduce latency.
- High Dimensionality: Post-hoc explanations may not accurately reflect model reasoning.
- Explanation Fidelity: Packet sequences and flow dynamics require specialized handling.

1.1.3. Operational Challenges

- Dataset Bias: Benchmark datasets may not reflect real-world traffic distributions.
- Feature Engineering: Packet-level vs. flow-level features require different approaches.

- Deployment Constraints: Production IDS must balance detection performance with throughput.

2 Review Methodology

This section presents the organized identification, screening, eligibility evaluation, and synthesis of studies on feature extraction-feature importance, and explainable AI (XAI) techniques in network intrusion detection systems (NIDS), following PRISMA-inspired guidelines.

2.1 Identification of Studies

To conduct a comprehensive literature survey, we have searched the targeted publications between 2020 and 2025 using predefined keyword combinations related to Network traffic classification, intrusion detection, feature extraction, feature engineering, and explainability through the open access repository and major academic databases like SpringerLink, ScienceDirect, IEEE Xplore, ACM Digital Library, MPDI, Wiley Online Library, Hindawi, arXiv, and Frontiers.

The initial search retrieved approximately 150 records, covering Feature extraction and selection methods for NIDS, deep learning-based intrusion detection models, Explainable AI and feature importance techniques, Systematic and narrative reviews, and emerging IoT and IIoT-focused intrusion detection studies.

2.2 Screening

During the screening process, redundant records across databases were removed, resulting in approximately 125 unique studies. During title and abstract screening, articles were excluded if they focused exclusively on signature-based or rule-based IDS, addressed non-network intrusion domains, and lacked any discussion of feature extraction, feature selection, or interpretability. After screening, approximately 85 studies remained for further evaluation.

2.3 Eligibility Assessment

A Comprehensive analysis was done on the remaining studies to evaluate Methodological clarity, Precise handling of feature extraction, feature selection, or feature importance, use of standard NIDS datasets, relevance to ML/DL-based intrusion detection, and availability of open-access full text.

We have excluded the publications that lacked relevance to feature engineering or explainability and studies that did not provide adequate methodological aspect. Finally, 38 studies were selected for our review.

2.4 Included Studies and Classification

The final 38 studies were classified into five thematic and conceptual categories, which enables systematic review and comparative analysis.

2.4.1. Feature Extraction–Centric Studies

Several studies highlighted the importance of feature extraction in detection performance, employing statistical, flow-based, spatiotemporal, and deep learning-based extraction techniques. Yu *et al.*⁽⁴⁾, Mao *et al.*⁽¹⁰⁾, Wang *et al.*⁽²²⁾, and Lypa *et al.*⁽²³⁾ demonstrated that multi-level and spatiotemporal feature extraction improves attack differentiation, particularly for massive and temporal attacks.

In IoT and large-scale traffic scenarios, Deep learning-based feature extraction using CNNs, RNNs, Auto-encoders, and attention mechanisms confirmed better performance than manual feature engineering^{(10), (12), (24)}.

2.4.2. Feature Selection and Optimization Studies

Feature selection techniques—including Information Gain, correlation-based filtering, permutation importance, and hybrid selection strategies—were widely used to reduce dimensionality and computational cost. The Authors, Kasongo and Sun⁽⁵⁾, recognized the significance of feature selection on UNSW-NB15, while more recent studies focused on intelligent and optimized selection mechanisms^{(6), (13)}. Several studies highlighted that feature selection combined with an ensemble model or hybrid classifiers notably improves detection accuracy while maintaining scalability^{(11), (14), (15)}.

2.4.3. Explainable AI and Feature Importance Studies

Explainability comes into sight as a major research trend post-2022. To explain ML and DL-based intrusion detection models, SHAP and LIME were applied by Khan⁽¹⁷⁾, Wei *et al.*⁽¹⁸⁾, Sharma *et al.*⁽¹⁹⁾, Gaspar *et al.*⁽²⁰⁾, and Islam *et al.*⁽²¹⁾.

XAI techniques were used to identify globally influential traffic features, provide class-wise explanations for attacks such as DDoS, and enhance trust and transparency in IoT and IIoT environments⁽²⁵⁾.

However, most studies applied XAI after the training phase, without interpreting explanation outputs for systematic feature refinement.

2.4.4. Hybrid and Ensemble-Based Approaches

Hybrid architectures combining feature extraction, feature selection, and ensemble learning were explored to balance performance and robustness. Ensemble models with optimized feature subsets proposed by Doost *et al.*⁽¹⁴⁾, Abdelaziz *et al.*⁽¹⁵⁾, and Ford and Berry⁽²⁶⁾ outperform single classifiers.

While these approach shown the benefits of combining multiple feature perspectives but often lacked interpretability or addressing redundant and highly correlated feature analysis.

2.4.5. Systematic Reviews and Comparative Studies

Broad overviews of ML/DL-based NIDS are reviewed by Zhang *et al.*⁽¹⁾, Ramya Chinnasamy *et al.*⁽²⁾, Sharma and Arora⁽³⁾, and Almuhanha and Dardouri⁽²⁷⁾. These studies have focused on model architectures, data imbalance issues, and performance metrics.

The Present review demonstrated the benefits of explicit integration of XAI with correlation analysis and feature dependency assessment, which was largely absent in existing studies. Table 1 shows the comparative study of existing work with their key contribution and gap exist in their work.

3 Overview

This study has reviewed papers published between 2020 and 2025, focusing on data imbalance, deep learning architecture, and spatiotemporal feature extraction challenges in machine learning/ Deep learning based approach for Network Intrusion Detection Systems. While these reviews provide valuable comparisons, the exhibits some limitations. First, feature importance and explainability are often treated as secondary considerations, rather than as essential components of the NIDS pipeline. Second, correlation analysis among netflow features remains underexplored, regardless of the facts that many NIDS datasets show evidence of high multicollinearity. Third, recent and realistic datasets—such as NF-ToN-IoT-v2- are largely missing from prior reviews. Much of the literature continues to rely on older benchmarks, limiting the generalizability of reported findings to modern-day network environments.

3.1 Datasets And Evaluation Metrics

3.1.1. Common datasets

NSL-KDD and CICIDS variants have been used for transparent IDS evaluations; CIC-IDS2018 was used in a DNN feature-selection comparison⁽⁴⁾; UNSW-NB15 and ToN-IoT are used for diverse IDS tasks and IIoT evaluations⁽²⁹⁾,⁽¹⁶⁾.

3.1.2. Metrics

Accuracy and F1-score are commonly reported; some works include explanation fidelity and qualitative expert assessments for XAI outputs.

3.1.3. Dataset issues

Class imbalance is explicitly handled via oversampling in some studies to improve DNN training⁽³³⁾; NetFlow conversions and format differences impact available features and require careful preprocessing⁽¹⁶⁾.

3.2 Feature Importance Techniques

This section summarizes Feature Importance Techniques such as model-specific, model-agnostic, information-theoretic, and statistical approaches used in traffic classification and compares their empirical strengths and weaknesses. Figure 1 shows the evolution of feature importance in network security.

Table 1. Comparative study of existing work

Ref.	Year	Feature Handling Technique	ML / DL Model	Key Contribution	Research Gap / Limitation
(1)	2025	Spatiotemporal features	DL models	Review of DL-based IDS	Does not cover XAI advances
(2)	2025	DL feature extraction	DL models	Systematic DL IDS review	No discussion on feature importance
(3)	2025	ML/DL features	ML & DL	Anomaly IDS review	Lacks XAI discussion
(4)	2024	Feature extraction	ML models	Improved IDS via features	No interpretability
(5)	2020	Feature selection	ML models	Early FS impact study	Outdated dataset
(6)	2024	Intelligent FS	ML models	Performance optimization	No XAI
(7)	2024	Feature embedding	Stacking ML	IDS for imbalanced data	No interpretability
(8)	2023	FS vs extraction	ML models	Comparative FS study	No interpretability
(9)	2025	Attention-based features	DAMLAN	Adaptive IDS	Attention not interpretable
(10)	2025	Multi-level extraction	Inductive network	Hierarchical feature learning	No explainability
(11)	2025	Feature selection + sensitivity	Hybrid ML	Ransomware detection	Narrow attack scope
(12)	2025	Autoencoder features	Autoencoder	Efficient IoT IDS	Latent features unexplained
(13)	2025	Feature selection	ML models	Optimized feature subset	Ignores explainability
(14)	2025	Feature selection + ensemble	Ensemble ML	Improved detection	Lacks interpretability analysis
(15)	2025	Permutation FI	Random Forest	Feature-guided RF IDS	Ignores feature correlation effects
(16)	2022	Statistical feature analysis	RF, SVM	Empirical feature relevance study	No explainability; ignores feature redundancy
(17)	2024	Feature importance analysis	ML models	Transparent IDS using XAI	Lacks correlation-aware feature pruning
(18)	2023	Flow-based features	ML classifiers	Class-wise DDoS explanations	Limited to single attack type
(19)	2023	Automatic feature learning	Deep NN	Explainable IoT IDS	No correlation analysis of learned features
(20)	2024	Engineered features	MLP	LIME vs SHAP comparison	Does not link explanations to feature reduction
(21)	2023	Feature attribution	ML/DL	Trustworthy IDS explanations	Lacks dataset-level feature correlation
(22)	2024	Spatio-temporal extraction	DL model	Temporal modeling	Black-box nature
(23)	2025	Tool-based extraction	ML models	Feature extraction tools comparison	No importance ranking
(24)	2025	CNN feature extraction	CNN	Explainable CNN IDS	Correlation among CNN features ignored
(28)	2024	Temporal features	LSTM	Pattern-based explanation	No global feature stability analysis
(29)	2023	Feature extraction + ensemble	DL ensemble	Explainable IIoT IDS	High computational overhead
(25)	2025	Feature fusion	Double-layer GRU	Interpretable GRU IDS	No evaluation on modern IoT datasets
(26)	2025	Feature standardization	Dynamic ensemble	Robust ensemble model	Feature importance not analyzed
(27)	2025	Learned features	Hybrid ML/DL	Hybrid anomaly detection	No explainability
(30)	2025	Statistical features	Decision Tree	Lightweight IDS	Low detection accuracy
(31)	2025	Contrastive features	Bayesian GMM	Self-supervised IDS	Feature semantics unclear
(32)	2024	Optimized feature selection	DNN	Performance-driven feature selection	Black-box DL without interpretability
(33)	2025	Selected traffic features	ML models	Real-time IoT IDS	Feature selection not justified
(34)	2025	Feature selection theory	ML models	FS foundations	No practical XAI integration
(35)	2024	Feature selection	ML models	FS comparison	No explainability
(36)	2023	FS + DL	DL models	Data-driven IDS	No explanation
(37)	2024	Feature selection	ML models	Improved IoT IDS	No XAI
(38)	2023	Feature extraction	ML models	ML-based IoT IDS	Outdated attack patterns

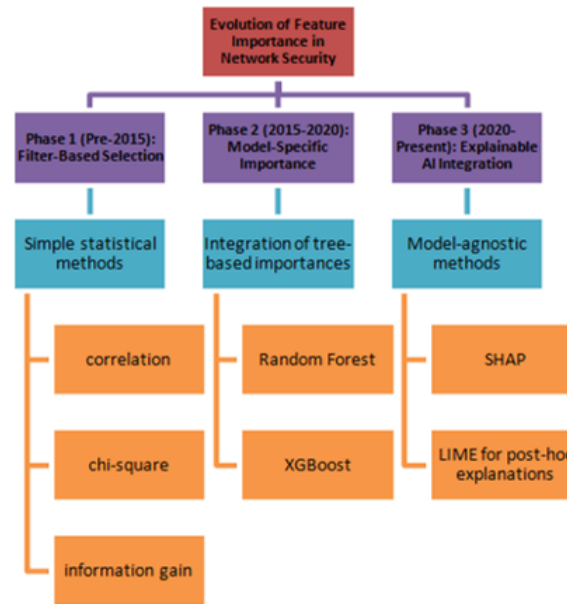


Fig 1. Evolution of Feature Importance in Network Security

Comparative study of common feature importance techniques is shown in Table 2. Feature selection and explanation methods are described below.

3.2.1. Model-specific methods

Model-specific method integrates importances based on trees (Random Forest, XGBoost). It assigns importance via model internals such as tree split gains or learned attention; e.g., XGBoost-derived importances were used to preselect top-20 flow features for an MLP DDoS classifier⁽¹⁸⁾. It has black-box model limitations in security contexts.

3.2.2. Model-agnostic explainers

Model-agnostic explainers like SHAP and LIME are widely applied to produce global and local attributions for both traditional ML and deep models. Hybrid pipelines can combine selection and explanation. It focuses on operational trust and regulatory compliance.

3.2.3. Information-theoretic scores

Information-theoretic scores such as information gain have been used as filter selectors and compared against correlation and chi-square methods in multi-dataset studies.

3.2.4. Statistical tests and Correlation analysis

Statistical tests and correlation analyses focus on dimensionality reduction for computational efficiency. These are applied as fast filters (correlation-based FS reported best results in one DNN intrusion study)⁽⁴⁾. This method has limited consideration of model interpretability.

3.3 Application To Attack Detection

Feature importance methods have been applied across DDoS, intrusion, IoT/IIoT attacks, and malware-like traffic, with flow-level and packet-level features both targeted, depending on attack modality and model choice.

3.3.1. DDoS detection

Packet- and flow-level DDoS classifiers using LSTM and MLP models have been explained using SHAP and SHAPPD, which revealed packet fields and flow statistics driving predictions^{(18), (28)}.

Table 2. Comparative study of common feature importance techniques

Method	Type	Mechanism	Advantages	Limitations	Observations
Tree-Based Importance	Model-Specific	Measures feature contribution via split information gain or impurity reduction	Fast computation, directly available from trained models, intuitive interpretation	Biased toward high-cardinality features, sensitive to tree structure	XGBoost-derived importances used to preselect top-20 flow features for MLP DDoS classifier, achieving >99% accuracy ⁽¹⁸⁾ Random Forest importances for initial feature ranking before deep model training ⁽³³⁾
Neural Network Attention Mechanisms		Learned weights indicating input importance at each layer	Built into model architecture, provides layer-wise insights	Requires specialized architectures, may not reflect global importance	Attention-based LSTM models for packet sequence classification ⁽²⁸⁾
SHAP (SHapley Additive exPlanations)	Model-Agnostic	Game-theoretic approach Computing each feature's marginal contribution	Consistent, theoretically grounded, supports global and local explanations	Computationally expensive for large models ($O(2^n)$ for exact computation)	Global feature importance for DDoS detection models ⁽¹⁸⁾ Local explanations for individual attack classification decisions SHAPPD Extension: Dependency-aware SHAP for LSTM-based DDoS detection, outperforming vanilla SHAP in capturing temporal patterns ⁽²⁸⁾
LIME (Local Interpretable Model-agnostic Explanations)		Fits interpretable surrogate models locally around predictions	Fast computation, intuitive linear approximations	Sensitive to locality definition and perturbation strategy	Explaining DNN/CNN intrusion detection decisions ^{(19), (29)} Real-time explanation generation for operational IDS ⁽²⁰⁾
Permutation Importance		Measures performance drop when feature values are randomly shuffled	Model-agnostic, captures feature interactions	Computationally expensive, requires multiple model evaluations	Limited evidence in reviewed corpus (insufficient documentation)
Information Gain / Mutual Information	Information-Theoretic Approaches	Measures entropy reduction or mutual dependence between features and target	Fast filter method, no model training required	Univariate view, misses feature interactions	Used across UNSW-NB15, CIC-IDS2018, ToN-IoT for ranking features ⁽¹⁶⁾
Chi-Square Test		Statistical test of independence between categorical features and target	Simple, efficient for categorical data	Requires discretization for continuous features	Compared against correlation methods in multi-dataset studies ⁽¹⁶⁾
Correlation-Based Feature Selection		Identifies features highly correlated with target and lowly correlated with each other	Reduces redundancy, computationally efficient	Linear relationships only, misses nonlinear dependencies	Best performer in DNN study on CIC-IDS2018: F1 = 99.93% with 57/79 features ⁽³³⁾ Outperformed SHAP and information-gain-based selection in direct comparison
ANOVA F-Test		Variance-based statistical test for feature relevance	Well-established, handles continuous features	Assumes normal distribution, univariate analysis	Used as baseline filter in comparative studies

3.3.2. General intrusion detection

DNNs/CNNs and traditional ML applied to NSL-KDD, UNSW-NB15, and CIC datasets have relied on filter selection and XAI to improve IDS interpretability and performance.

3.3.3. IIoT and IoT scenarios

Ensemble DL with SHAP/LIME explanations was evaluated on ToN_IoT for industrial IDS tasks^{(16), (29)}.

3.4 Practical Analysis

3.4.1. Correlation Analysis

We have applied the spearman correlation method on the NF-TON_IOT_v2 dataset, which consists of 43 features. Figure 2 shows the correlation analysis of the dataset. In our existing study⁽³⁰⁾, we have got 99.31% accuracy for Binary classification, where we first used Spearman correlation with a value of 0.7. Then, based on spearman correlation, we dropped 25 features as shown in the Table 3.

In a direct comparison on CIC-IDS2018, Correlation-Based Feature Selection outperformed SHAP and information-gain-based selection for a DNN, achieving F1 = 99.93% with 57 features⁽³²⁾. SHAP extensions that model temporal or dependency structure (SHAPPD) have been proposed to better explain sequential models like LSTM, outperforming vanilla SHAP in justification richness for packet-sequence decisions⁽⁸⁾. Several works combine filter selection with SHAP/LIME post-hoc explanations to balance scalability and interpretability, but standardized evaluation of explanation fidelity is still limited in the IDS literature^{(17), (21)}.

Table 3. Dropped features based on Spearman (0.7)

Sr. No.	Dropped Features
1	'MIN_IP_PKT_LEN',
2	'DST_TO_SRC_SECOND_BYTES',
3	'SERVER_TCP_FLAGS',
4	'OUT_PKTS',
5	'ICMP_IPV4_TYPE',
6	'MAX_IP_PKT_LEN',
7	'MIN_TTL',
8	'ICMP_TYPE',
9	'DURATION_IN',
10	'DST_TO_SRC_AVG_THROUGHPUT',
11	'DNS_QUERY_ID',
12	'NUM_PKTS_UP_TO_128_BYTES',
13	'TCP_WIN_MAX_IN',
14	'SRC_TO_DST_SECOND_BYTES',
15	'TCP_FLAGS',
16	'TCP_WIN_MAX_OUT',
17	'RETRANSMITTED_IN_PKTS',
18	'OUT_BYTES',
19	'DURATION_OUT',
20	'IN_PKTS',
21	'CLIENT_TCP_FLAGS',
22	'RETRANSMITTED_OUT_PKTS',
23	'DNS_QUERY_TYPE',
24	'MAX_TTL',
25	'LONGEST_FLOW_PKT']

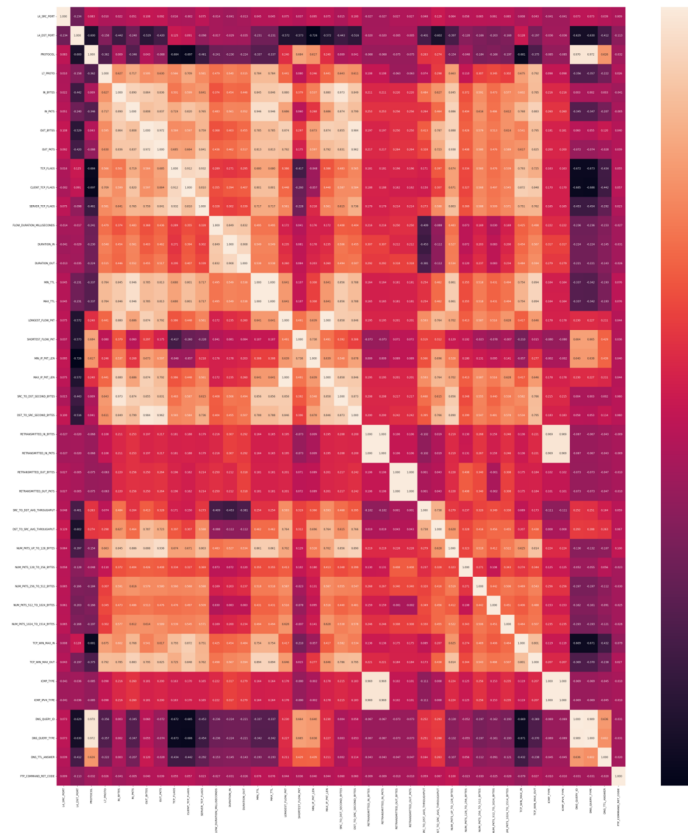


Fig 2. Spearman Correlation Analysis

4 Conclusion

This study gives a comprehensive and timely synthesis of advancements in NIDS with a focused prominence on the emerging convergence of explainable AI (XAI) and feature correlation analysis review for the researchers. Feature importance techniques have developed from simple statistical filters to sophisticated model-agnostic explainability methods, enabling both high-performance attack detection and transparent, auditable decision-making. The convergence of machine learning, explainable AI, and network security has produced practical systems achieving 99%+ accuracy with interpretability, deployed in resource-constrained environments, and trusted by security analysts.

However, significant challenges remain: standardizing explanation evaluation, ensuring cross-dataset generalization, bridging the deployment gap between research and practice, and maintaining robustness against adversarial attacks. Addressing these challenges requires interdisciplinary collaboration among machine learning researchers, cyber-security experts, human-computer interaction specialists, and industry practitioners.

References

- 1) Zhang Y, Muniyandi RC, Qamar F. A Review of Deep Learning Applications in Intrusion Detection Systems: Overcoming Challenges in Spatiotemporal Feature Extraction and Data Imbalance. *Appl Sci*. 2025;15:1552. <https://doi.org/10.3390/app15031552>.
- 2) Chinnasamy R, Subramanian M, Easwaramoorthy SV, Cho J. Deep learning-driven methods for network-based intrusion detection systems: A systematic review. *ICT Express*. 2025;11(1):181–215. <https://doi.org/10.1016/j.icte.2025.01.005>.
- 3) Sharma N, Arora B. Machine Learning and Deep Learning Models for Anomaly Intrusion Detection in Networks: A Systematic Review. *SN Comput Sci*. 2025;6:832. <https://doi.org/10.1007/s42979-025-04352-z>.
- 4) Yu X, Huang Y, Zhang Y, Song M, Jia Z. Network Intrusion Traffic Detection Based on Feature Extraction. *Computers, Materials & Continua*. 2024;78(1):473–492. <https://doi.org/10.32604/cmc.2023.044999>.
- 5) Kasongo SM, Sun Y. Performance analysis of intrusion detection systems using a feature selection method on the UNSW-NB15 dataset. *Journal of Big Data*. 2020;7:105. <https://doi.org/10.1186/s40537-020-00379-6>.
- 6) Aljehane NO, Mengash HA, Hassine SBH, Alotaibi FA, Salama AS, Abdelbagi S. Optimizing intrusion detection using intelligent feature selection with machine learning model. *Alexandria Engineering Journal*. 2024;91:39–49. <https://doi.org/10.1016/j.aej.2024.01.073>.

- 7) Talukder MA, Islam MM, Uddin MA, Hasan KF, Sharmin S, Alyami SA, et al. Machine learning-based network intrusion detection for big and imbalanced data using oversampling, stacking feature embedding and feature extraction. *Journal of Big Data*. 2024;11:33. <https://doi.org/10.1186/s40537-024-00886-w>.
- 8) Ngo VD, Vuong TC, Luong TV, Tran H. Machine learning-based intrusion detection: feature selection versus feature extraction. arXiv. 2023. Available from: <https://arxiv.org/abs/2307.01570>.
- 9) Alrayes FS, Amin SU, Hakami NA, Alzaylaee MK, Kashmeery T. Intrusion detection model on network data with deep adaptive multi-layer attention network (DAMLAN). *Computer Modeling in Engineering & Sciences*. 2025;144(1):581–614. <https://doi.org/10.32604/cmescs.2025.065188>.
- 10) Mao J, Yang X, Hu B, Lu Y, Yin G. Intrusion detection system based on multi-level feature extraction and inductive network. *Electronics*. 2025;14(1):189. <https://doi.org/10.3390/electronics14010189>.
- 11) Zhang K, Wang Y, Bhatti UA, et al. Enhanced ransomware attacks detection using feature selection, sensitivity analysis, and optimized hybrid model. *Journal of Big Data*. 2025;12:245. <https://doi.org/10.1186/s40537-025-01289-1>.
- 12) Feng Y, Sakurai K. An efficient detection mechanism of network intrusions in IoT environments using autoencoder and data partitioning. *Computers*. 2025;13(10):269. <https://doi.org/10.3390/computers13100269>.
- 13) Alzubi QM, Makhadmeh SN, Sanjalawe Y. Optimizing Intrusion Detection: Advanced Feature Selection and Machine Learning Techniques Using the CSE-CIC-IDS2018 Dataset. *Journal of Advances in Information Technology*. 2025;16(3). Available from: <https://www.jait.us/articles/2025/JAIT-V16N3-283.pdf>.
- 14) Doost PA, Moghadam SS, Khezri E, et al. A new intrusion detection method using ensemble classification and feature selection. *Sci Rep*. 2025;15:13642. <https://doi.org/10.1038/s41598-025-98604-w>.
- 15) Abdelaziz MT, Radwan A, Mamdouh H, et al. Enhancing network threat detection with random forest-based NIDS and permutation feature importance. *Journal of Network and Systems Management*. 2025;33(2). <https://doi.org/10.1007/s10922-024-09874-0>.
- 16) Sarhan M, Layeghy S, Portmann M. Feature Analysis for Machine Learning-based IoT Intrusion Detection. 2022. <https://doi.org/10.21203/rs.3.rs-2035633/v1>.
- 17) Khan MA. Explainable AI and Machine Learning Models for Transparent and Scalable Intrusion Detection Systems. *Journal of Information Systems Engineering and Management*. 2024 Dec;9(4s):1576–1588. <https://doi.org/10.52783/jisem.v9i4s.12115>.
- 18) Wei Y, Jang JS, Singh A, Sabrina F, Camtepe S. Classification and Explanation of Distributed Denial-of-Service (DDoS) Attack Detection using Machine Learning and Shapley Additive Explanation (SHAP) Methods;vol. abs/2306.17190. arXiv.org. 2023. <https://doi.org/10.48550/arXiv.2306.17190>.
- 19) Sharma B, Sharma LK, Lal C. Explainable artificial intelligence for intrusion detection in IoT networks: A deep learning based approach. *Expert Systems With Applications*. 2023 Sept;238:121751–121751. <https://doi.org/10.1016/j.eswa.2023.121751>.
- 20) Gaspar D, Silva P, Silva C. Explainable AI for Intrusion Detection Systems: LIME and SHAP Applicability on Multi-Layer Perceptron. *IEEE Access*. 2024;p. 1–1. <https://doi.org/10.1109/ACCESS.2024.3368377>.
- 21) Islam TT, Syfullah MK, Rashed MG, Das D. Bridging the Gap: Advancing the Transparency and Trustworthiness of Network Intrusion Detection with Explainable AI. 2023. <https://doi.org/10.21203/rs.3.rs-3263546/v1>.
- 22) Wang C, Zhang W, Hao H, Shi H. Network traffic classification model based on spatio-temporal feature extraction. *Electronics*. 2024;13(7):1236. <https://doi.org/10.3390/electronics13071236>.
- 23) Lypa B, Horyn I, Zagorodna N, Tymoshchuk D, Lechachenko T. Comparison of feature extraction tools for network traffic data. arXiv. 2025. Available from: <https://arxiv.org/abs/2501.13004>.
- 24) Ebrahimi F, Javidan R, Akbari R, et al. Intrusion detection in the internet of things using convolutional neural networks: an explainable AI approach. *Cybersecurity*. 2025;8:66. <https://doi.org/10.1186/s42400-025-00369-2>.
- 25) Wijaya MR, Hanafi M. Towards Interpretable Intrusion Detection: A Double-Layer GRU with Feature Fusion Explained by SHAP and LIME. *Informatik Journal*. 2025;21(3). <https://doi.org/10.52958/iftk.v21i3.12187>.
- 26) Ford J, Berry HS. Advancing Network Threat Detection Through Standardized Feature Extraction and Dynamic Ensemble Learning. Concord, NC, USA. 2025. <https://doi.org/10.1109/SoutheastCon56624.2025.10971597>.
- 27) Almuhamma R, Dardouri S. A deep learning/machine learning approach for anomaly based network intrusion detection. *Front Artif Intell*. 2025;8:1625891. <https://doi.org/10.3389/frai.2025.1625891>.
- 28) AsSadhan B, Bashaiwath A, Binsalleeh H. Enhancing Explanation of LSTM-based DDoS Attack Classification using SHAP with Pattern Dependency. *IEEE Access*. 2024 Jan;p. 1–1. <https://doi.org/10.1109/access.2024.3421299>.
- 29) Shtayat MM, Hasan KM, Sulaiman R, Islam S, Rehman AU. An Explainable Ensemble Deep Learning Approach for Intrusion Detection in Industrial Internet of Things. *IEEE Access*. 2023. <https://doi.org/10.1109/access.2023.3323573>.
- 30) Bhosle SN, Pagare J. Detecting Network Attacks Using Decision Tree Classifier;vol. 2822. Cham. Springer. 2026. https://doi.org/10.1007/978-3-032-17843-5_15.
- 31) Liu L, Xu M. A network intrusion detection method based on contrastive learning and Bayesian Gaussian mixture model. *Cybersecurity*. 2025;8:59. <https://doi.org/10.1186/s42400-025-00364-7>.
- 32) Anagha AS, Thomas C, Sreelatha G. Enhanced Intrusion Detection Through Optimized Feature Selection in Deep Neural Networks. Kottayam, India. 2024. <https://doi.org/10.1109/ISENSE63713.2024.10872230>.
- 33) Ahmed AA. Enhancing IoT security: A machine learning-based intrusion detection system for real-time threat detection and mitigation. *Journal of Education and Science*. 2025;34(4):45–61. <https://doi.org/10.33899/jes.v34i4.49257>.
- 34) Hailes S, Westphal C, Musolesi M. Feature selection for network intrusion detection. ACM. 2025. <https://doi.org/10.1145/3690624.3709339>.
- 35) Westphal C, Hailes S, Musolesi M. Feature selection for network intrusion detection. arXiv. 2024. Available from: <https://arxiv.org/abs/2411.11603>.
- 36) Zhang L, Liu K, Xie X, Bai W, Wu B, Dong P. A data-driven network intrusion detection system using feature selection and deep learning. *Journal of Information Security and Applications*. 2023;78:103606. <https://doi.org/10.1016/j.jisa.2023.103606>.
- 37) Almohaimeed M, Albalwy F. Enhancing IoT Network Security Using Feature Selection for Intrusion Detection Systems. *Applied Sciences*. 2024;14(24):11966. <https://doi.org/10.3390/app142411966>.
- 38) Musleh D, Alotaibi M, Alhaidari F, Rahman A, Mohammad RM. Intrusion Detection System Using Feature Extraction with Machine Learning Algorithms in IoT. *Journal of Sensor and Actuator Networks*. 2023;12(2):29. <https://doi.org/10.3390/jsan12020029>.