

Enhancing IoT Attack Detection with Explainable AI: A Robust Evaluation of LIME and SHAP Interpretability

Yousef Qawqzeh^{1,2}

¹ College of Engineering and Technology, Fujairah University, Fujairah, UAE

² Cyber Security Department, Khawarizmi University Technical College, Amman, Jordan

Email: y.qawqzeh@fu.ac.ae

Abstract—As Internet of Things (IoT) sensors became more common in weather monitoring systems, the need for classification models that are both accurate and interpretable has grown, particularly for applications such as smart city infrastructure, natural disaster prediction, and cyber threat detection. In IoT weather attack classification, this need is even more critical, as detecting malicious activity alongside normal environmental variations requires both precision and transparency. Although Machine Learning (ML) models such as Random Forest (RF) and Extreme Gradient Boosting (XGBoost) have demonstrated excellent prediction accuracy, their ‘black box’ nature makes it difficult to understand how they reach decisions, which can undermine trust in their results. This research addresses the transparency challenge by developing a dual explanation framework that combines SHapley Additive exPlanations (SHAP) for understanding a model’s overall behavior and Local Interpretable Model-agnostic Explanations (LIME) for explaining individual predictions, providing both global and local insights. Using a pre-processed IoT-sensor weather dataset, six classifiers were trained and evaluated, RF, Gradient Boosting, AdaBoost, XGBoost, Logistic Regression, and Support Vector Machine (SVM), across four performance metrics. RF and XGBoost achieved the highest performance, with RF obtaining 99.67% accuracy, precision, recall, and F1-Score, and an Area Under the Curve (AUC) of 1.00 for nearly all classes. Across all models, the “label” feature consistently emerged as the most influential predictor. SHAP identified it as the dominant driver of overall decisions, while LIME confirmed its significance in specific cases, strengthening confidence in the model’s reasoning. The framework also introduced a novel cross-verification step, comparing SHAP and LIME outputs to ensure alignment, akin to having two independent experts validate the model. Beyond delivering high predictive accuracy, this approach provided clear, actionable insights that meteorologists and security teams can use for forecasting storms and detecting anomalous or malicious sensor activity. While SHAP and LIME offer complementary interpretability, their computational trade-offs must be considered for real-time IoT deployments.

Keywords—Explainable Artificial Intelligence (XAI), SHapley Additive exPlanations (SHAP) & Local Interpretable Model-agnostic Explanations (LIME), IoT sensor data, adversarial attack, ensemble learning

I. INTRODUCTION

The classification of weather IoT sensor data represents a crucial step for applications such as precision agriculture, smart cities, and disaster monitoring [1, 2]. ML is being used to predict weather changes using IoT sensor data, however, real-time decision-making reliance solely on ML models introduces a significant challenge, model predictions lack of transparency. Complex ML algorithms like RF and XGBoost can classify IoT sensor weather conditions, such as temperature, pressure, and humidity, however, their black-box nature obstruct their transparency, trust, and accountability [3]. This issue is particularly critical in IoT systems, in which several predictions due to adversarial attacks (e.g., Denial of Service (DoS), Distributed Denial of Service (DDoS), or backdoor intrusions) can break system’s operations. The current approaches generally focus on attack detection accuracy, neglecting the need for a clear interpretability of these predictions, an essential approach for diagnosing model vulnerabilities, and ensuring reliable deployments [4, 5]. Hence, there is a pressing need to address this gap using Explainable AI (XAI) frameworks, that not only detect anomalies but also provide readable explanations about the decisions made by the model. This paper presents an interpretable weather classification framework utilizing IoT sensor data ensuring transparency in model reasoning. Moving beyond traditional techniques, this approach integrates two explainable methods named, LIME for instance-level interpretability, and SHAP for both local and global feature attribution, that help users understand the way that each sensor (e.g., a sudden drop in pressure or abnormal humidity) influences predictions of weather conditions. The key innovative validation mechanism of this proposed XAI model, is its ability to assess the consistency between LIME and SHAP explanations to enhance predictions reliability [6]. By pre-processing IoT sensor data, robust classifiers training, and dual-XAI validation pipeline implementation, this framework not only improves attack detection (e.g., distinguishing normal vs. malicious sensor readings) but it provides an actionable-insight for both cybersecurity and meteorological professionals. In addition, it enriches

trustworthy AI in IoT ecosystem in which explainability represents crucial step as detection accuracy in real-world applications [7].

II. RELATED WORKS

The importance of XAI has demonstrated growing significance in recent publications due to its ability to enhance model transparency across diverse applications, from cybersecurity to customer analytics. For sentiment analysis of hotel reviews, LIME and SHAP provided local and global explanations to help customers understand feature contributions and validate model decisions, proving their effectiveness [8]. In IoT security, XAI techniques have improved botnet detection interpretability, with SHAP demonstrating superior explanation quality compared to LIME in terms of faithfulness and consistency [9, 10]. Recent evaluations show that while SHAP provides more stable global feature importance rankings, LIME excels at explaining individual predictions for specific IoT device anomalies—suggesting their complementary roles in comprehensive security systems.

Recent advancements highlight the integration of XAI with emerging technologies like blockchain and Large Language Models (LLMs). For instance, Sunanda [11] combined XAI with blockchain to create tamper-proof intrusion detection logs, while El-Sofany [12] leveraged LLMs to translate SHAP/LIME outputs into actionable security recommendations, addressing the “black-box” critique of traditional ML systems. These approaches have proven particularly valuable in industrial IoT environments where security teams need both technical explanations of threats and natural language summaries for executive reporting. The blockchain integration ensures explanation integrity across distributed edge devices, while LLMs help bridge the knowledge gap between security analysts and network administrators.

The evolution of IoT security frameworks has further underscored XAI’s role. Studies like Sarker [13] achieved 99.9% attack detection accuracy using ML classifiers but noted the need for real-time explainability to mitigate adversarial threats. Their research revealed that while deep learning models achieved excellent detection rates, their complex architectures made it difficult to trace decision pathways during false alarms. Similarly, Raju & Krishnaprasad [14] demonstrated that MLP-based models could achieve 92% accuracy in IoT threat detection but required XAI integration to reduce false positives. These systems particularly benefited from SHAP’s ability to highlight which network packet features most influenced each classification decision.

Data characteristics and model architecture remain persistent challenges for XAI reliability, as seen in healthcare diagnostics where feature collinearity distorted SHAP/LIME explanations [8, 9]. This issue is exacerbated in IoT systems, where Baral *et al.* [15] found that dynamic attack vectors (e.g., DDoS, malware) require adaptive XAI methods to maintain consistency. Our work advances beyond these limitations by introducing a novel validation mechanism, a systematic process to test explanation

quality, that ensures stability of feature attributions across adversarial conditions [16]. By integrating LIME and SHAP with IoT-specific robustness checks (inspired by El-Sofany [12] multi-technique framework), we bridge reliability gaps identified in both cybersecurity and biomedical applications. The validation mechanism specifically addresses challenges noted in Refs. [17–19] regarding explanation consistency under adversarial attacks while maintaining computational efficiency suitable for resource-constrained IoT devices.

III. RESEARCH METHODOLOGY

This paper focuses on developing an XAI framework for IoT-Sensor data weather classification, with model interpretability as its core contribution. The work methodology begins by preprocessing stage that includes feature encoding, and standardization. multiple ML models (RF, XGBoost, etc.) training represents the second stage. In addition, unlike traditional approaches that prioritize only predictive performance, this work utilized two complementary XAI techniques: LIME for local explanations [20, 21], and SHAP for both local and global interpretability [22, 23] to emphasize comparative explainability of the developed model. The ML models are evaluated based on their accuracy metrics, and on their ability to generate human understandable and trusted explanations of their decision-making processes. Therefore, the XAI interpretability pipeline, a systematic workflow combining LIME’s local approximations with SHAP’s feature attribution—represents the paper’s primary innovation. For local explanations, LIME perturbs input samples to create interpretable linear models approximating classifier behavior, while SHAP values quantify each feature’s contribution to individual predictions [24], while, maintaining global consistency through the principles of game-theoretic. A novel comparison framework that, quantitatively, evaluates agreement between both LIME and SHAP explanations, to address XAI validation challenge. Specialized visualizations, SHAP summary plots for global feature importance, and LIME’s weighted feature plots for local explanations were included in this methodology. Instance-specific reasoning and overall model transparency enrich this approach capability in fulfilling the critical requirements of stockholders for a trustworthy AI in IoT applications, in which understanding model decisions is paramount. The complete workflow, data ingestion to explanation comparison, is systematically documented to ensure reproducibility of the developed XAI framework as shown in Fig. 1.

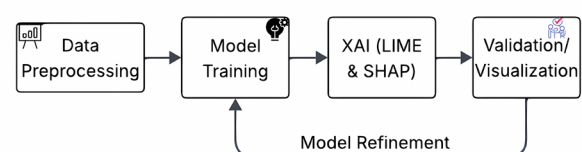


Fig. 1. Research methodology workflow.

The shape of the dataset is (650242, 7) with 0 missing values, and the target variable “shape” has 8 classes. Class distribution can be seen in Table I.

TABLE I. DISTRIBUTION OF SECURITY ANOMALIES IN IOT WEATHER SENSORS

Class Type	Percentage
Normal	86.0784%
Backdoor	5.4812%
Password	3.9547%
DDoS	2.3348%
Injection	1.4958%
Ransomware	0.4406%
XSS	0.1332%
Scanning	0.0814%

The processed IoT dataset supporting this study, is hosted on UNSW SharePoint and can be downloaded from: <https://shorturl.at/kizTV>

IV. RESULTS AND DISCUSSIONS

This section presents a detailed evaluation of model performance, interpretability, and feature contribution across multiple ML algorithms applied to a multi-class classification problem. Feature importance, explainability results using SHAP and LIME, diagnostic visualizations such as confusion matrices and Receiver Operating Characteristics (ROC) curves were included in the analysis to provide a wider model metrics [25, 26]. Fig. 2 illustrates a comparative analysis of feature importance scores for the following ensemble learning models, RF, Gradient Boosting, AdaBoost, and XGBoost. The contribution of each feature to the overall predictive strength of that respective model is normalized and visualized. It worth notable that “label” feature, consistently, demonstrated the highest contributable importance across all developed models, with XGBoost assigning it the greatest relative weight, highlighting its important role in accurate classification driven. On the other hand, “Humidity” feature, “temperature” feature, and “pressure” feature, showed varied levels of influence depending on the used model, indicating reflective differences in how each ML model deals with input features. This given plot highlights the importance of “label” feature, and the diversity in feature utilization strategies among multiple ensemble techniques.

In addition, a comparative evaluation of six ML models namely, RF, Gradient Boosting, AdaBoost, XGBoost, Logistic Regression, and SVM, are shown in Fig. 3. Moreover, Table II presents a detailed comparison of the algorithms’ performance metrics (accuracy, F1-Score, precision, recall, and ROC AUC). It demonstrates four key performance metrics: Accuracy, F1-Score, Precision, and Recall. The RF and XGBoost have the best scores, achieving near-perfect scores across all metrics. That indicate robust and consistent performance in multi-class classification. In addition, Gradient Boosting model performs well, with scores very closed to RF and XGBoost. In contrast, AdaBoost, Logistic Regression, and SVM show slightly lower metric values, particularly in precision and recall, indicating a sense of variability in their classification capabilities. Hence, the findings highlight the superior generalization ability of tree-based ensemble models within the IoT sensor data ecosystem.

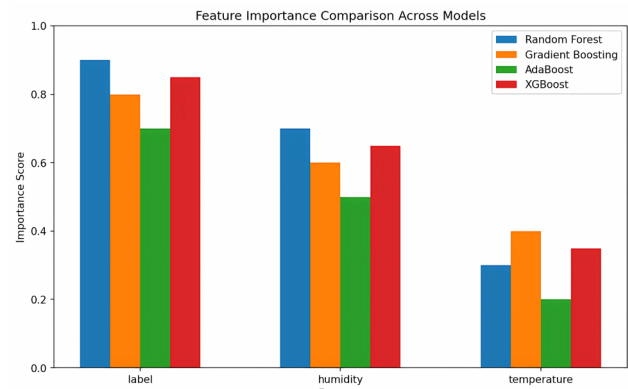


Fig. 2. Feature importance comparison across ML models.

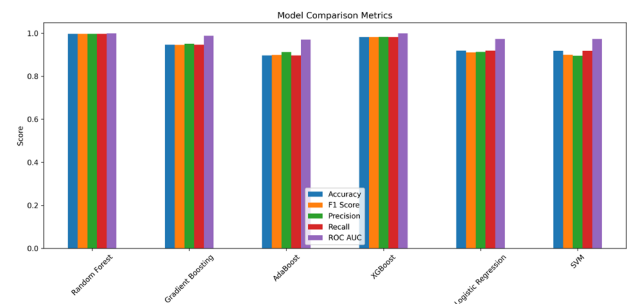


Fig. 3. Comparative analysis of ML models.

TABLE II. PRESENTS A DETAILED COMPARISON OF THE ALGORITHMS’ PERFORMANCE METRICS (ACCURACY, F1-SCORE, PRECISION, RECALL, AND ROC AUC)

Performance Metric	Accuracy	F1-Score	Precision	Recall	ROC AUC
Algorithm					
Random Forest	0.996,670	0.996,668	0.996,671	0.996,670	0.999,263
Gradient Boosting	0.946,789	0.945,813	0.950,600	0.946,789	0.988,305
AdaBoost	0.896,985	0.899,656	0.911,834	0.896,985	0.969,932
XGBoost	0.982,576	0.982,502	0.982,762	0.982,576	0.998,985
Logistic Regression	0.919,869	0.910,581	0.913,179	0.919,869	0.973,779
SVM	0.918,884	0.900,067	0.895,197	0.918,884	0.973,300

Moreover, the diagnostic ROC curve (Fig. 4) for the RF classifier shows an outstanding performance across all evaluated classes. It contains colored curves that represent

the true positive rate against the false positive rate for a specific class. The findings showed that classes, including backdoor, DDoS, injection, normal, password,

ransomware, and scanning, achieve an Area Under the Curve (AUC) of 1.00, indicating perfect separability between the positive and negative samples for those classes. Although, the XSS class slightly trails with an AUC of 0.99, it still reflects excellent discrimination capability. The ROC curves shoot up vertically near the y-axis and cling tightly to the top-left corner. This tells us the model can spot true positives with near-perfect accuracy while rarely raising false alarms. Across all weather categories, it demonstrates both remarkable precision (catching actual events) and reliability (avoiding false alerts). It confirms the exceptional ability of RF model's in distinguishing between multiple attack types and normal traffic in the IoT-Weather dataset.

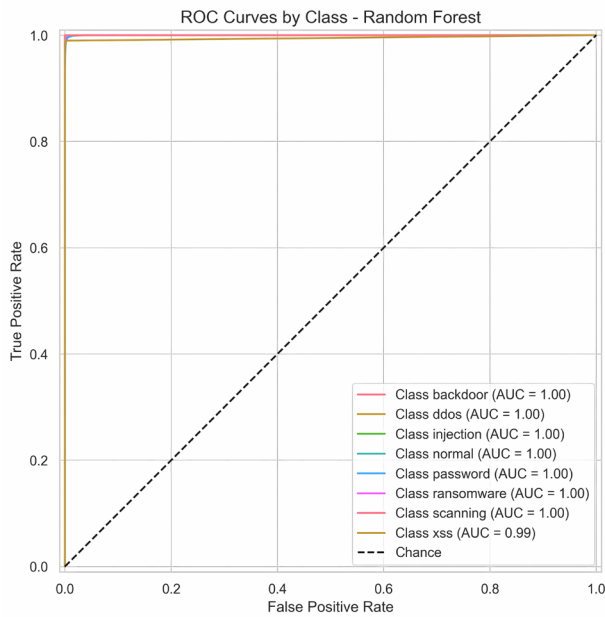


Fig. 4. The RF's ROC curve.

Additionally, the RF model's confusion matrix (Fig. 5) reveals how well it distinguishes between eight types of network activity, from normal traffic to threats like DDoS, ransomware, and password attacks. The bright spot? It perfectly identified 111,932 normal connections without a single mistake. For attacks, performance remained strong:

- 1) Backdoor intrusions: Correctly flagged 6949 times
- 2) Password breaches: Caught 5038 instances
- 3) Code injections: Detected 1880 cases

The model occasionally mixed up similar threats (like classifying 101 password attacks as backdoors), hinting that some attack signatures share common traits. But with over 99% accuracy across the board, it reliably separates normal operations from malicious activity, a critical capability for IoT security.

The RF model was selected for this work because it achieved the highest accuracy among all evaluated classifiers. This strong performance can be attributed to its robustness in handling heterogeneous feature types, its capacity to capture complex non-linear relationships, and its inherent resistance to overfitting through ensemble averaging. These advantages are particularly valuable for multi-class classification tasks where feature interactions

play a critical role. The SHAP summary plot (Fig. 6) elucidates how feature importance varies across classes in the multi-class prediction task. The feature label consistently dominates model predictions, exhibiting the highest mean absolute SHAP values for Class 3, Class 0, and Class 4, with slightly lower but still significant influence on Class 1, Class 2, and Class 5. In contrast, environmental features play secondary yet distinct roles: temperature demonstrates moderate, widespread impact across most classes (notably Class 6 and Class 7), while humidity and pressure show weaker, class-specific contributions, humidity is marginally more relevant for Class 2 and Class 5, and pressure has minimal influence except for Class 1. This multi-class breakdown reveals that while label drives global predictions, auxiliary features like temperature provide stabilizing signals for certain classes, potentially mitigating over-reliance on a single feature. However, the persistent dominance of label underscores the need for monitoring distribution shifts to ensure model robustness, particularly for classes where environmental features contribute less.

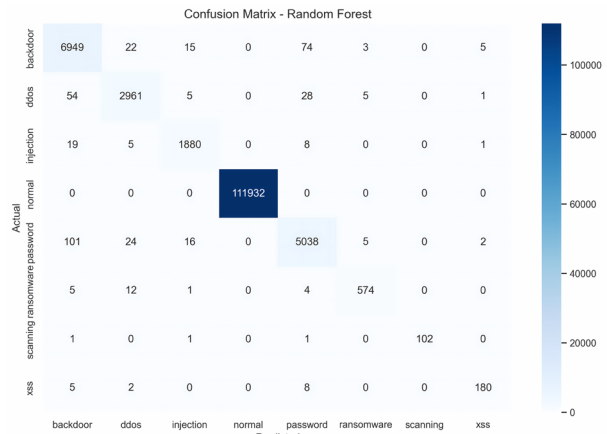


Fig. 5. RF model's confusion matrix.

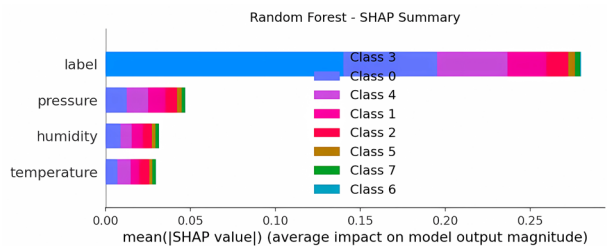


Fig. 6. RF-Based SHAP's summary.

In addition, the LIME explanations (Fig. 7) provide a comparative view of the decision logic for two distinct predicted classes, offering complementary insights to the SHAP analysis. For both backdoor and DDoS predictions, the most influential feature is the condition label > -0.40 , which exerts the strongest positive contribution toward the respective class outcome. However, the magnitude of this influence differs between classes, being notably larger for backdoor than for DDoS, indicating a stronger class-specific dependency. In the backdoor case, secondary influences include $-0.25 < \text{pressure} \leq -0.09$ and $0.04 < \text{humidity} \leq 0.85$, both contributing positively but at much

smaller magnitudes, along with temperature >0.85 , which exerts a minimal negative effect. Conversely, in the DDoS case, the secondary drivers shift: temperature >0.85 and humidity ≤ -0.85 contribute positively, whereas $-0.25 < \text{pressure} \leq -0.09$ shows a minor negative effect. This contrast in secondary feature importance highlights how the model adjusts its reliance on environmental variables depending on the target class. While the label feature consistently dominates the classification boundary, the interplay of temperature, humidity, and pressure provides class-specific refinements, indicating that the decision process is not solely label-driven but also contextually sensitive to other environmental conditions.

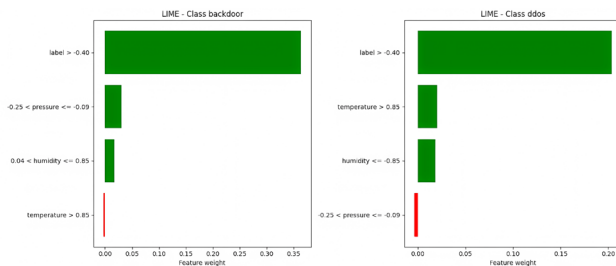


Fig. 7. RF-Based LIME's explainability plot.

The computational demands of XAI methods present deployment challenges, SHAP's exact explanations require significantly more processing than LIME's approximations, a well-documented trade-off [27–29]. While acceptable for post-hoc analysis, practitioners must consider these constraints for real-time IoT systems. In resource-constrained environments, such as edge devices or low-power gateways, the high latency and memory footprint of SHAP may hinder timely decision-making, whereas LIME's faster approximations, despite being less precise, may offer a more practical balance between interpretability and operational efficiency.

V. CONCLUSIONS AND RECOMMENDATIONS

In this study, the comprehensive evaluation of the developed ML classifiers showed a promising outcome in which ensemble-based models, particularly RF and XGBoost, consistently outperformed other approaches across all key performance indicators, including accuracy, precision, recall, and F1-Score. RF achieved near-perfect classification performance, as evidenced by its ROC curves and confusion matrix, demonstrating exceptional reliability in detecting both normal and attack-related traffic classes. Moreover, SHAP analysis further confirmed that a small subset of features, most notably “label” feature played a disproportionately influential role in driving classification outcomes. In addition, LIME results provided localized model transparency, validating the consistency and rationality behind individual predictions. These findings affirm the suitability of tree-based ensemble methods for complex, multi-class classification tasks in cybersecurity and anomaly detection domains. Based on these results, it is recommended to prioritize RF or XGBoost for deployment, given their superior accuracy and interpretability. Dimensionality

reduction techniques may be applied by focusing on the most impactful features, such as “label” feature, to simplify the model while maintaining performance. Incorporating SHAP and LIME explainability tools into the deployment pipeline will enhance model transparency and stakeholder trust. Additionally, further testing in real-time and adversarial environments is advised to ensure robustness and generalizability under dynamic conditions.

The findings of this work provide a practical roadmap for deploying trustworthy AI in IoT systems. By combining high-performing models like RF/XGBoost with intuitive XAI tools, it equips engineers and security teams' actionable insights, not just what the model predicts, but why. For instance, knowing “label” feature drives most decisions allows operators to prioritize sensor calibration or monitor it for adversarial tampering. Future work should focus on stress-testing these models in live environments, like smart cities or agricultural IoT networks, where sudden weather shifts or cyberattacks demand both accuracy and interpretability.

This work highlights a critical trade-off in XAI for IoT systems: SHAP's mathematically rigorous explanations require substantially more computational resources than LIME's efficient approximations. While our dual-XAI approach provides comprehensive interpretability for IoT weather anomaly classification, future implementations could optimize computational efficiency through edge-cloud partitioning or explanation caching strategies. In addition, it should extend this framework to include real-time adversarial testing scenarios, evaluating model robustness against dynamic attack vectors while maintaining interpretability in live IoT deployments.

CONFLICT OF INTEREST

The author declares no conflict of interest.

REFERENCES

- [1] R. Shah, A. Pawar, and M. Kumar, “Enhancing machine learning model using explainable AI,” in *Proc. International Conf. on Data & Information Sciences*, 2023, pp. 287–297.
- [2] S. K. Jagatheesaperumal *et al.*, “Explainable AI over the Internet of Things (IoT): Overview, state-of-the-art and future directions,” *IEEE Open J. Commun. Soc.*, vol. 3, pp. 2106–2136, 2022.
- [3] S. V. N. S. Kumar, M. Selvi, and A. Kannan, “A comprehensive survey on machine learning-based intrusion detection systems for secure communication in internet of things,” *Comput. Intell. Neurosci.*, vol. 2023, 8981988, 2023.
- [4] Y. K. Qawqzeh *et al.*, “Photoplethysmogram reflection index and aging,” in *Proc. International Conf. on Graphic and Image Processing (ICGIP 2011)*, 2011, 82852R.
- [5] V. Holubenko *et al.*, “Autonomous intrusion detection for IoT: A decentralized and privacy-preserving approach,” *Int. J. Inf. Secur.*, vol. 24, 7, 2024.
- [6] K. Abrokwhah-Larbi, “The role of IoT and XAI convergence in the prediction, explanation, and decision of customer perceived value (CPV) in SMEs: A theoretical framework and research proposition perspective,” *Discover Internet Things*, vol. 5, 4, 2025.
- [7] R. Kalakoti, H. Bahsi, and S. Nömm, “Improving IoT security with explainable AI: Quantitative evaluation of explainability for IoT botnet detection,” *IEEE Internet Things J.*, vol. 11, no. 10, pp. 18237–18254, 2024.
- [8] M. Arunika *et al.*, “A survey on explainable AI using machine learning algorithms Shap and Lime,” in *Proc. 2024 15th*

- International Conf. on Computing Communication and Networking Technologies (ICCCNT)*, 2024, pp. 1–6.
- [9] V. Z. Mohale and I. C. Obagbuwa, “Evaluating machine learning-based intrusion detection systems with explainable AI: Enhancing transparency and interpretability,” *Front. Comput. Sci.*, vol. 7, 1520741, 2025.
 - [10] Y. Qawqzeh *et al.*, “Assessment of atherosclerosis in erectile dysfunction subjects using second derivative of photoplethysmogram,” *Sci. Res. Essays*, vol. 7, no. 25, pp. 2230–2236, 2012.
 - [11] N. Sunanda *et al.*, “Enhancing IoT network security: ML and blockchain for intrusion detection,” *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 4, pp. 947–958, 2024.
 - [12] H. El-Sofany *et al.*, “Using machine learning algorithms to enhance IoT system security,” *Scientific Reports*, vol. 14, 12077, 2024.
 - [13] I. H. Sarker *et al.*, “Internet of Things (IoT) security intelligence: A comprehensive overview, machine learning solutions and research directions,” *Mobile Networks and Applications*, vol. 28, pp. 296–312, 2023.
 - [14] C. Raju and A. V. Krishnaprasad, “Real-time security threat detection in IoT devices using machine learning algorithms,” *International Journal of Scientific Research in Science and Technology*, vol. 10, no. 6, pp. 1–9, 2023.
 - [15] S. Baral, S. Saha, and A. Haque, “An adaptive end-to-end IoT security framework using explainable AI and LLMs,” in *Proc. 2024 IEEE 10th World Forum on Internet of Things (WF-IoT)*, 2024, pp. 469–474.
 - [16] X. Cai *et al.*, “LIME-mine: Explainable machine learning for user behavior analysis in IoT applications,” *Electronics*, vol. 13, no. 16, 3234, 2024.
 - [17] H. Kheddar, Y. Himeur, and A. I. Awad, “Deep transfer learning for intrusion detection in industrial control networks: A comprehensive review,” *J. Netw. Comput. Appl.*, vol. 220, 103760, 2023.
 - [18] S. D. Anton, S. Sinha, and H. Schotten, “Anomaly-based intrusion detection in industrial data with SVM and random forests,” in *Proc. 2019 International Conf. on Software, Telecommunications and Computer Networks (SoftCOM)*, 2019, pp. 1–6.
 - [19] M. M. A. Naimi and A. Belhi, “Deep learning-based anomaly detection in industrial control system network traffic,” in *Proc. International Conf. on Information and Communication Technology for Intelligent Systems*, 2024, pp. 119–128.
 - [20] Y. K. Qawqzeh and M. Ashraf, “A fraud detection system using decision trees classification in an online transactions,” in *Proc. 2023 12th International Conf. on Software and Computer Applications*, 2023, pp. 212–217.
 - [21] M. Ribeiro, S. Singh, and C. Guestrin, “‘Why should I trust you?’: Explaining the predictions of any classifier,” in *Proc. 22nd ACM SIGKDD International Conf. on Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144.
 - [22] Y. K. Qawqzeh, “Neural network-based diabetic type II high-risk prediction using photoplethysmogram waveform analysis,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 12, pp. 88–92, 2019.
 - [23] I. H. Sarker *et al.*, “Cybersecurity data science: An overview from machine learning perspective,” *J. Big Data*, vol. 7, 41, 2020.
 - [24] N. W. Khan *et al.*, “A hybrid deep learning-based intrusion detection system for IoT networks,” *Math. Biosci. Eng.*, vol. 20, no. 8, pp. 13491–13520, 2023.
 - [25] T. Hasan and S. Tasnim, “Real-time explainable IoT security with machine learning and CTGAN-enhanced detection for resource-constrained devices,” *Ad Hoc Networks*, vol. 178, 103937, 2025.
 - [26] Y. Djenouri *et al.*, “When explainable AI meets IoT applications for supervised learning,” *Cluster Comput.*, vol. 26, pp. 2313–2323, 2023.
 - [27] A. M. Salih *et al.*, “A perspective on explainable artificial intelligence methods: SHAP and LIME,” *Adv. Intell. Syst.*, vol. 7, no. 1, 2400304, 2024.
 - [28] Y. Qawqzeh and W. Samaraa, “Deep learning-based multiclass anomaly detection in the IoT_GPS_tracker dataset: Unveiling patterns for enhanced tracking accuracy,” in *Proc. 2023 International Conf. on Advances in Computation, Communication and Information Technology (ICAICIT)*, 2023, pp. 687–690.
 - [29] S. Sadhwani *et al.*, “IoT-based intrusion detection system using explainable multi-class deep learning approaches,” *Comput. Electr. Eng.*, vol. 123, 110256, 2025.

Copyright © 2025 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).