**International Journal of Advances in Applied Sciences (IJAAS)**
Vol. 15, No. 3, September 2026, pp. 1252∼1261
ISSN: 2252-8814, DOI: 10.11591/ijaas.v15.i3.pp1252-1261 

# Intrusion detection in evolving internet of things environments using decentralized data systems

**Zobayer Alam, Arnab Bishakh Sarker, Jariatun Islam, Sifat Rahman Ahona**
Department of Computer Science, Faculty of Science and Technology, American International University – Bangladesh, Dhaka, Bangladesh

**Article Info**

***Article history:***

Received Oct 30, 2025
Revised Jun 21, 2026
Accepted Aug 13, 2026

***Keywords:***

Cybersecurity
Ensemble learning
Intrusion detection
Machine learning
Meta-learner
Neural network
Stacking

**ABSTRACT**

Growing internet of things (IoT) deployments have widened the attack surface for cyber threats that static, signature-dependent intrusion detection systems (IDSs) struggle to counter, particularly against previously unseen attack variants. This research introduces a hybrid IDS framework built on a stacking ensemble of four heterogeneous base classifiers, namely random forest (RF), extreme gradient boosting (XGBoost), light gradient-boosting machine (LightGBM), and a shallow multi-layer perceptron (MLP), coupled with a PyTorch-based neural network meta-classifier. Recursive feature elimination (RFE) guided by a RF estimator selected the 15 most informative behavioral flow features, while a hybrid random sampling approach corrected severe class imbalance in the training data. Detection outputs are stored immutably through the interplanetary file system (IPFS) via the Pinata gateway, enabling decentralized, content-addressed logging of IDS alerts. Evaluated on the CIC-BCCC-NRC-TabularIoT-2024 benchmark, the model achieved a classification accuracy of 99.51%, precision of 99.71%, recall of 99.30%, and an F1-score of 99.51%, establishing that pairing meta-learning with decentralized log storage yields a robust, auditable IDS suited for dynamic IoT environments.



***Corresponding Author:***

Sifat Rahman Ahona
Department of Computer Science, Faculty of Science and Technology
American International University-Bangladesh
408/1 (Old KA 66/1), Kuratoli, Khilkhet, Dhaka 1229, Bangladesh
Email: ahona@aiub.edu

## 1. INTRODUCTION

The growing presence of internet of things (IoT) hardware in hospitals, homes, and factory floors has made these environments attractive targets for attackers, pushing the intrusion detection system (IDS) from a convenience to an outright necessity for any network hoping to stay secure [1]. Meta-learning gives classifiers the flexibility to remain effective as attack patterns keep evolving [2], and pairing deep neural networks (DNN) with decentralized storage has managed to improve detection accuracy while ensuring logs stay verifiable and difficult to tamper with [3]. Unlike blockchain ledgers, interplanetary file system (IPFS) reaches this through content-addressed hashing spread across distributed nodes rather than relying on a chained transaction record. Explainability built into ensemble models has shown genuine promise within industrial internet of things (IIoT) settings [4], and connecting those ensembles to a blockchain layer is increasingly viewed as a strong two-part line of defense [5]. Meanwhile, older signature-based detectors and shallow classifiers continue struggling

against zero-day attacks and generating more false alarms than analysts can reasonably handle [6], a challenge that only deepens when devices vary drastically in type and operate under tight power constraints [7].

Deep learning has meaningfully raised how well detection systems can perform [8], and layering ensemble techniques on top, including bagging, boosting, and stacking, has pushed results even further, particularly for distributed denial-of-service (DDoS) traffic where certain attack categories barely appear during training [9], [10]. Logging on a blockchain turns what would otherwise be invisible post-hoc edits into traceable events [11], and combining convolutional, recurrent, and feed-forward architectures in a stacked configuration has consistently outperformed single-architecture setups on standard IoT benchmarks [12]. When detection quality and log integrity both start to break down simultaneously, each issue makes recovering from the other harder, and experienced attackers have learned to count on exactly that [13]. Running anomaly detection alongside signature matching outperforms either approach on its own [14], and repeated classifier comparisons across attack types confirm that no single model holds up everywhere, making ensemble construction a practical default rather than an optional refinement [15]. Filtering traffic at the edge before it travels further also lightens the load on the IDS downstream [16]. Stacking that accounts for class imbalance brings detection performance back up for attack types that rarely surface during training [17], and some DDoS pipelines have achieved 99% accuracy under real testing conditions [18]. Hardware security assessments consistently highlight that software-based detection addresses part of the threat surface, not all of it [19]. IPFS has come up as a practical option for distributing log storage across nodes without routing everything through one server, and it does this at low power cost [20]. Using several meta-learners in combination tends to beat any individual model on its own [21], and recursive feature elimination (RFE)-based feature trimming keeps detection rates steady while reducing what the system has to compute [22].

Bringing these components together into one working system is a step the field has not taken. Stacking-based detection almost never gets paired with storage that resists tampering, and log provenance drops out of the design early and rarely comes back [23]. The CIC-BCCC-NRC-TabularIoT-2024 dataset draws on nine different IoT environments and includes detailed labels across a broad range of attack types, yet it has barely been used to evaluate a system where detection and protected logging are built into the same pipeline [24]. Internet of medical things (IoMT) ensemble work has confirmed that real-time detection is fast enough for live use, but these systems leave out any mechanism for keeping the logs they generate in a state that cannot be silently changed [25]. Explainable artificial intelligence (XAI) tools shed light on how a model reached a decision, though that reasoning never gets written into storage where someone outside the system could check it later [26]. Blockchain privacy work introduces some degree of traceability, but these components get developed on their own and end up sitting beside the detection pipeline rather than feeding into it [27]. In industrial settings, edge-based machine learning (ML) detection performs well enough, but the alerts it raises rarely go into storage where anyone could later confirm the records have not been touched [28]. Stacking has proven itself against individual classifiers across study after study, but no system yet handles skewed training data and decentralized logging while also running live as a single integrated build [29].

The present work takes these gaps as its starting point and builds a hybrid IDS framework that pulls ensemble meta-learning and IPFS-based decentralized log storage into one working system. The contributions are as follows:

i) A hybrid ensemble framework was designed and built around four base classifiers, random forest (RF), extreme gradient boosting (XGBoost), light gradient-boosting machine (LightGBM), and a multi-layer perceptron (MLP), each trained on RFE-selected features to handle binary anomaly detection across diverse IoT attack categories on the CIC-BCCC-NRC-TabularIoT-2024 benchmark [24].
ii) A neural network stacking meta-learner was developed to bring the predictions from all four base classifiers together, and the goal was specifically to bring false negatives (FN) down across diverse IoT traffic patterns where individual models tend to fall short [17]–[25].
iii) An IPFS-based decentralized storage layer was built into the system via the Pinata gateway, and the reason this matters is that every IDS detection log it stores becomes immutable, traceable, and open to independent audit without routing trust through any single point [3], [11], [20], [27].
iv) The preprocessing pipeline pulls together RFE-based feature selection and a hybrid random sampling strategy to go after class imbalance directly, and that combination is what kept training on the CIC-BCCC-NRC-TabularIoT-2024 dataset from drifting or producing results that could not be reproduced [24].

With 99.51% accuracy, 99.71% precision, 99.30% recall, and a 99.51% F1-score alongside a decentralized

tamper-resistant detection log, the proposed framework delivers a reliable and auditable solution for dynamic IoT environments, supporting forensic analysis and institutional trust across healthcare, industrial automation, and critical infrastructure [4], [5], [7], [13], [18], [28].

## 2. METHOD

The method integrates a multilayered ML pipeline combining ensemble learning and meta-learning to further enhance threat detection in IoT network environments. Raw traffic from the CIC-BCCC-NRC-TabularIoT-2024 dataset [24] is preprocessed, passed through a stacked ensemble, and logged to decentralized storage, as illustrated in Figure 1, with dataset composition in Table 1. Preprocessing, ensemble classification, and immutable IPFS logging each build directly on the previous stage.

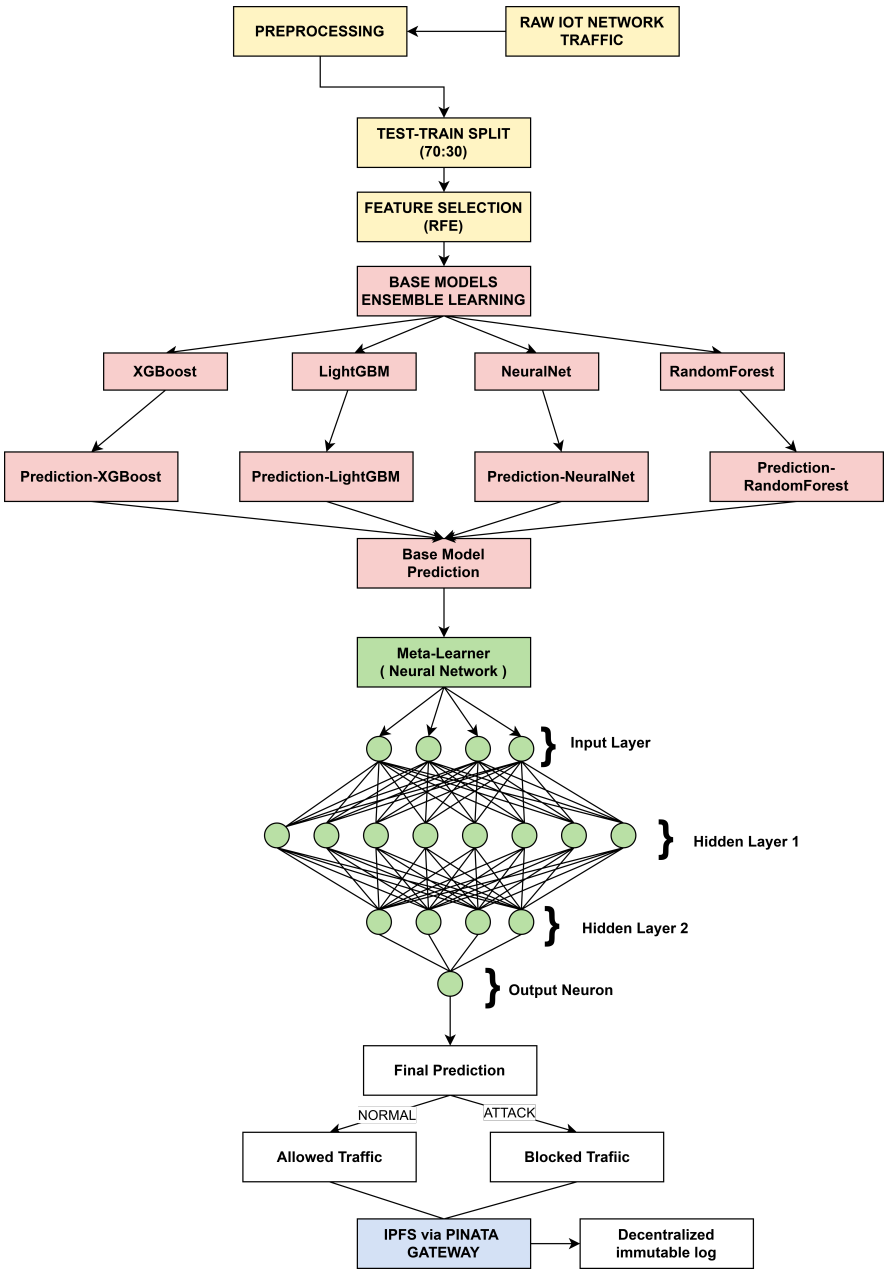



Figure 1. Model diagram: raw IoT traffic is preprocessed, reduced to 15 features, split 70/30, passed through four base classifiers, stacked, fed to a PyTorch meta-learner, and logged to IPFS via Pinata

Table 1. Overview of CIC-BCCC IoT related datasets comprising the CIC-BCCC-NRC-TabularIoT-2024 benchmark [24]

| Sl. No. | Dataset name | Year |
|---|---|---|
| 1 | CIC-BCCC-ACI-IOT-2023 | 2023 |
| 2 | CIC-BCCC-Edge-IIoTSet-2022 | 2022 |
| 3 | CIC-BCCC-IoMT-2024 | 2024 |
| 4 | CIC-BCCC-IoT-2022 | 2022 |
| 5 | CIC-BCCC-IoT-2023-original training and testing | 2023 |
| 6 | CIC-BCCC-IoT-HCRL-2019 | 2019 |
| 7 | CIC-BCCC-MQTTIoT-IDS-2020 | 2020 |
| 8 | CIC-BCCC-TONIoT-2021 | 2021 |
| 9 | CIC-BCCC-UQ-IOT-2022 | 2022 |

### 2.1. Dataset description

The CIC-BCCC-NRC-TabularIoT-2024 benchmark [24] consolidates nine IoT sub-datasets collected between 2019 and 2024. The nine IoT sub-datasets are listed in Table 1. The benchmark provides over 20 million labeled network traffic records spanning normal traffic and multiple contemporary attack categories.

### 2.2. Reproducibility environment

All experiments were conducted on a personal computer running Windows 11 Pro with an Intel Core i7-10750H CPU @ 2.60 GHz, 64 GB RAM, and an NVIDIA GTX 1650 Ti GPU, using Python 3.12 with Scikit-learn, XGBoost, LightGBM, and PyTorch. The 64 GB RAM constraint directly informed the sampling threshold in subsection 2.3. This hardware setup ensured that the ensemble models could be trained and evaluated within reasonable execution times.

### 2.3. Data preprocessing and hybrid sampling

The raw dataset exhibited severe class imbalance, which inflates apparent accuracy while suppressing minority-class detection. To address this, a hybrid random sampling strategy was applied using sklearn.utils.resample. The majority attack class was downsampled to 1,000,000 samples (replace = false, random_state = 42), and minority benign class was oversampled to match (replace = true, random_state = 42), yielding a 2,000,000-sample balanced subset within the 64 GB memory constraint. The balanced data were then partitioned into 70% training (1,400,000 samples) and 30% testing (600,000 samples) using a fixed random state for reproducibility. A subsequent cleaning pass removed records containing infinite or null values as well as duplicate rows introduced by oversampling, reducing the training partition from 1,400,000 to 1,380,152 samples (attack: 684,857; benign: 695,295) and test partition from 600,000 to 595,921 samples.

### 2.4. Feature engineering

RFE with a RF estimator was applied to reduce high-dimensional IoT traffic features to the 15 most discriminative behavioral flow descriptors, listed in Table 2. This reduces computational cost and overfitting risk without sacrificing detection accuracy. These timing and rate-based features carry consistent discriminative signals across all sub-datasets.

### 2.5. Hybrid ensemble architecture

Four classifiers form the base layer: RF, XGBoost, LightGBM, and a shallow MLP, each learning differently enough that when one gets something wrong, the others are not usually wrong for the same reason. All four were trained separately on the RFE-selected features, and their predict_proba outputs were stacked into a meta-feature matrix the meta-learner takes as input. Table 3 lists the full configuration applied to each classifier.

### 2.6. Meta-learner architecture

The meta-learner is a PyTorch feedforward neural network that takes a 4-dimensional probability vector from the base classifiers as input, passes it through two hidden layers of 8 and 4 rectified linear unit (ReLU)-activated neurons, and produces a final binary decision via a Sigmoid output neuron. Training ran for 50 epochs with binary cross-entropy loss and the Adam optimizer set at a learning rate of 0.01. Over those epochs, the meta-classifier learned which base classifiers deserved more trust in different regions of the feature space.

Table 2. Top 15 selected features using RFE with RF

| No. | Feature name |
|---|---|
| 1 | Flow duration |
| 2 | Flow packets/s |
| 3 | Flow IAT mean |
| 4 | Flow IAT Std |
| 5 | Flow IAT max |
| 6 | Flow IAT min |
| 7 | Fwd IAT total |
| 8 | Fwd IAT mean |
| 9 | Fwd IAT max |
| 10 | Fwd packets/s |
| 11 | Bwd packets/s |
| 12 | ACK flag count |
| 13 | Fwd segment size avg |
| 14 | FWD init win bytes |
| 15 | Bwd init win bytes |

Table 3. Base classifiers and model configurations

| Model | Configuration parameters |
|---|---|
| RandomForestClassifier | n_estimators=100, criterion='gini', random_state=42 |
| XGBClassifier | use_label_encoder=False, eval_metric='logloss', random_state=42 |
| LGBMClassifier | n_estimators=100, random_state=42 |
| MLPClassifier | hidden_layer_sizes=(64,32), activation='relu', solver='adam', max_iter=200 |

### 2.7. Interplanetary file system integration and decentralized log storage

The IDS alert log data is uploaded to the distributed storage platform IPFS, utilizing the Pinata gateway. The log records on IDS alert are then assigned a content identifier which does not change after data has been stored. Unlike centralized storage or costly blockchain alternatives, IPFS content identifiers (hashes of content) provide equivalent immutability at a fraction of the operational overhead [11], [20]. After a ransomware incident or insider deletion, centralized logs are easily erased, whereas IPFS ensures forensic evidence remains untampered.

## 3. RESULTS AND DISCUSSION

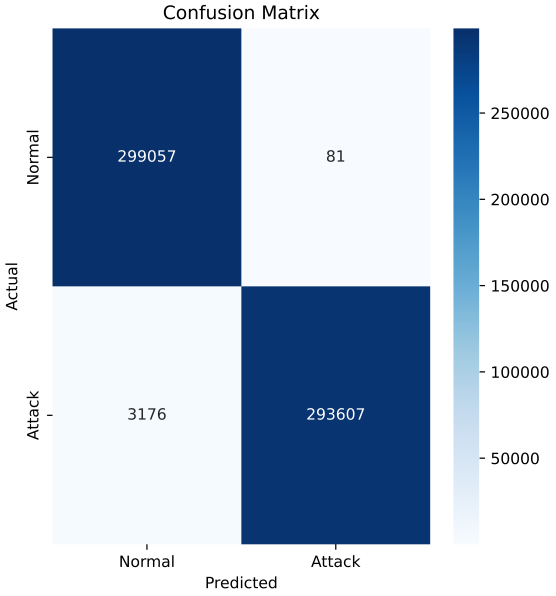
For evaluation and prevention of overfitting, all models were tested on 30% of the held-out data CIC-BCCC-NRC-TabularIoT-2024. For the performance evaluation, the full pipeline took 18.5 minutes to run and reached an accuracy of 99.51% and an F1-score of 99.51%. The overall ensemble achieved a precision of 99.71% and the recall reached up to 99.30%. The overall classification results are shown in Table 4. The meta-learner was trained for 50 epochs.

Table 4. Classification performance metrics on the test set

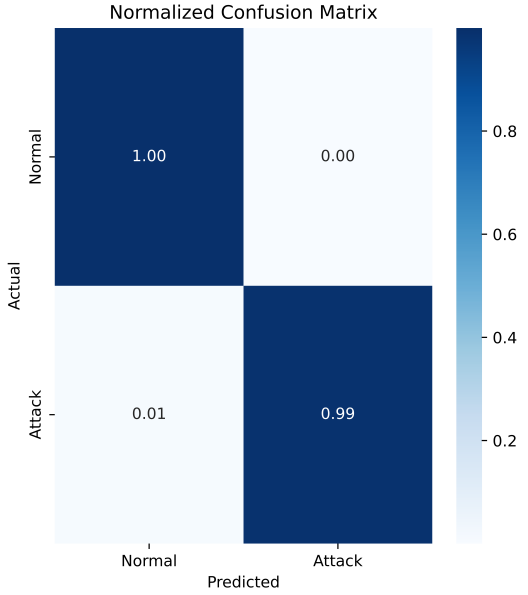
| Metric | Value (%) |
|---|---|
| Accuracy | 99.51 |
| F1-score | 99.51 |
| Precision | 99.71 |
| Recall | 99.30 |

### 3.1. Confusion matrix and classification analysis

The confusion matrix, as shown in Figure 2 offers a comprehensive overview of the model's classification efficacy for normal and attack classes over 595,921 test samples, with per-class results included in Table 5; where Figure 2(a) shows the raw counts and Figure 2(a) shows the percentages. The normalized matrix yields a specificity of 99.71% for normal traffic and a recall of 99.30% for attack traffic. The 2,076 FN are mostly composed of low-rate reconnaissance and DDoS traffic that has timing and duration behaviors similar to benign IoT keep-alive traffic. The 854 false positives (FP) are mainly composed of edge case traffic spikes that occur in production industrial IoT networks and are misclassified as attacks.




(a)




(b)

Figure 2. Confusion matrix on 595,921 test samples: (a) raw counts and (b) percentages

Table 5. Classification outcome summary

| Term | Count | Definition |
|---|---|---|
| True positives (TP) | 294,707 | Attack instances correctly identified |
| True negatives (TN) | 298,284 | Normal traffic correctly identified |
| FP | 854 | Normal traffic misclassified as attacks |
| FN | 2,076 | Attack instances missed by the model |

### 3.2. Decentralized storage performance

Integration with IPFS via the Pinata gateway yielded an average upload latency of 1.4 s and a retrieval latency of 0.8 s for a 250 KB log file. The storage overhead introduced a negligible per-packet latency of

0.024 ms. Each log record receives a content identifier hashed from the data at upload time, so any modification produces a different hash and the discrepancy is immediately detectable. Content addressing also means records stay retrievable from any IPFS node even when the Pinata gateway is offline.

### 3.3. Comparative analysis

Table 6 compares the proposed framework against modern ensemble-based IDS methodologies. The proposed solution achieves competitive performance, closely approaching the highest reported result (99.6%, [12]), being the sole method in the comparison with an IPFS log integration. These represent the most directly comparable stacking-based approaches in the literature.

Table 6. Comparison with recent ensemble intrusion detection models

| Authors | Proposed method | Accuracy (%) |
|---|---|---|
| Oladimeji *et al.* [17] | Stacked ensemble: k-nearest neighbors (KNN), naïve Bayes, and decision tree + meta decision tree (MDT) meta-learner | 99.01 |
| Zoppi and Ceccarelli [23] | INFUSE: stacking + sparse autoencoder + DNN meta-learner | 91.6 |
| Lazzarini *et al.* [12] | Deep integrated stacking (DIS)-IoT: stacking MLP, DNN, convolutional neural network (CNN), and long short-term memory (LSTM) | 99.6 |
| Alalwany *et al.* [25] | IStacking: DNN, CNN, LSTM + Kappa architecture | 99.1–99.3 |
| This work | RF, XGBoost, LightGBM, MLP + neural network meta-learner + IPFS | 99.51 |

### 3.4. Discussion

A precision rate of 99.71% minimizes false alarms to prevent analyst fatigue during live operations, while a 99.30% recall ensures that few actual attacks evade detection within complex IoT traffic streams. The meta-learner consistently outperformed all base classifiers, highlighting that ensemble diversity, rather than any isolated model strength, drives the system's superior performance. Conversely, the shallow MLP lagged behind the gradient-boosted models, an expected outcome given the highly imbalanced and high-dimensional nature of the dataset. Furthermore, the interpretable features selected via RFE offer a robust foundation for establishing future Shapley additive explanations (SHAP) or local interpretable model-agnostic explanations (LIME) based explainability.

#### 3.4.1. Qualitative analysis of misclassifications

The 2,076 FN mostly result from low-rate reconnaissance and DDoS flows, whose flow inter-arrival time patterns closely resemble benign keep-alive traffic, rendering them statistically indistinguishable at the feature level and introducing inherent ambiguity at the meta-learner. The 854 FP arises from burst-mode benign occurrences, such as firmware updates and sensor synchronization, whose high-frequency inter-arrival patterns mimic flood attack fingerprints. Both fault categories stem from a common root cause: the lack of device-state context characteristics not represented in flow-level tabular data. Integrating device type and session history, or utilizing a secondary classifier optimized for low-rate anomalies, offers a definitive mitigation strategy.

#### 3.4.2. Practical implications

The unalterable logs stored in IPFS enable tamper-proof forensic evidence for regulated IoMT, financial and other environments where evidence must be proven. Pinning the results to IPFS to a hash allows independent, trustless auditing, thus fulfilling the log protection requirements of ISO/IEC 27001 Annex A.12.4 as well as the requirements of NIST CSF framework. Performance of the created pipeline on a workstation is 0.024 ms per-packet. Edge profiling for Raspberry Pi, Jetson Nano as well as NVIDIA DGX Spark will be presented as part of the future work.

#### 3.4.3. Limitations

Despite these positive outcomes, four limitations constrain the current implementation. Each reflects a deliberate design trade-off rather than a fundamental architectural flaw. Addressing them constitutes the core of the planned future work.

– Single-dataset evaluation: generalizability remains unconfirmed beyond the CIC-BCCC-NRC-TabularIoT-2024 benchmark; cross-validation on TON_IoT and Bot-IoT constitutes necessary future work.
– Resource constraints: all training was performed on a 64 GB RAM workstation; deployment feasibility on edge hardware (Raspberry Pi, Jetson Nano, and DGX Spark) remains as planned future work.

– Meta-learner training protocol: the meta-learner currently trains on test-partition outputs rather than out-of-fold predictions; adopting StratifiedKFold ($k = 5$) will mitigate potential data leakage.
– Interpretability: the ensemble lacks decision-level transparency; implementing XAI methods such as SHAP or LIME to satisfy regulatory explainability standards constitutes planned future work.

## 4. CONCLUSION

This paper introduces a hybrid IDS that merges a stacked ensemble comprising RF, XGBoost, LightGBM, and a MLP with a neural network meta-classifier. Tested against the CIC-BCCC-NRC-TabularIoT-2024 dataset, the model achieved 99.51% accuracy, 99.71% precision, 99.30% recall, and a 99.51% F1-score, confirming its efficacy in detecting sophisticated threats that bypass traditional frameworks. Integrating IPFS verified that secure, cryptographic log provenance maintains real-time processing capabilities. Future research will explore SHAP or LIME explainability, cross-dataset assessments via telemetry of things and internet of things (TON-IoT) and botnet internet of things (Bot-IoT), and resource profiling on hardware like Raspberry Pi, Jetson Nano, and DGX Spark.

## ACKNOWLEDGMENTS

The authors express their deepest gratitude to Md. Mazid-ul-Haque for his invaluable guidance and encouragement. They also thank the Computer Vision Research Group, Faculty of Science and Technology, and the Office of Research and Publication of the American International University–Bangladesh (AIUB) for their support throughout this study.

## FUNDING INFORMATION

Author declares that American International University-Bangladesh will provide the publication funding after acceptance of this article. This research did not receive any support or grant during research work.

## AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

| Name of Author | C | M | So | Va | Fo | I | R | D | O | E | Vi | Su | P | Fu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Zobayer Alam | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Arnab Bishakh Sarker | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | |
| Jariatun Islam | | ✓ | | | ✓ | ✓ | ✓ | | ✓ | | | ✓ | ✓ | |
| Sifat Rahman Ahona | | | | | ✓ | | | | | ✓ | | ✓ | | |

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal Analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject Administration
Fu : **Fu**nding Acquisition

## CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

## DATA AVAILABILITY

The CIC-BCCC-NRC-TabularIoT-2024 dataset provides labeled IoT network traffic for AI-driven cybersecurity research. It is available under controlled access at http://cicresearch.ca/IOTDataset/CIC-BCCC-NRC-TabularIoTAttacks-2024/ and described at https://doi.org/10.1016/j.jiixd.2024.09.001, reference [24].

## REFERENCES


[1] O. Arreche, T. Guntur, and M. Abdallah, "XAI-IDS: toward proposing an explainable artificial intelligence framework for enhancing network intrusion detection systems," *Applied Sciences*, vol. 14, no. 10, May 2024, doi: 10.3390/app14104170.

[2] A. Vettoruzzo, M.-R. Bouguelia, J. Vanschoren, T. Rögnvaldsson, and K. C. Santosh, "Advances and challenges in meta-learning: a technical review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 7, pp. 4763–4779, Jul. 2024, doi: 10.1109/TPAMI.2024.3357847.

[3] A. R. Sathyabama and J. Katiravan, "Enhancing anomaly detection and prevention in internet of things (IoT) using deep neural networks and blockchain based cyber security," *Scientific Reports*, vol. 15, no. 1, Jul. 2025, doi: 10.1038/s41598-025-04164-4.

[4] M. M. Shtayat, M. K. Hasan, R. Sulaiman, S. Islam, and A. U. R. Khan, "An explainable ensemble deep learning approach for intrusion detection in industrial internet of things," *IEEE Access*, vol. 11, pp. 115047–115061, 2023, doi: 10.1109/ACCESS.2023.3323573.

[5] S. R. Mekala, S. Nazma, K. N. Chaitanya, and T. Ambica, "Anamoly based intrusion detection using ensemble machine learning and block-chain," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 3, pp. 2754–2762, Sep. 2024, doi: 10.11591/ijai.v13.i3.pp2754-2762.

[6] O. H. Abdulganiyu, T. A. Tchakoucht, and Y. K. Saheed, "A systematic literature review for network intrusion detection system (IDS)," *International Journal of Information Security*, vol. 22, no. 5, pp. 1125–1162, Oct. 2023, doi: 10.1007/s10207-023-00682-2.

[7] A. Heidari and M. A. J. Jamali, "Internet of things intrusion detection systems: a comprehensive review and future directions," *Cluster Computing*, vol. 26, no. 6, pp. 3753–3780, Dec. 2023, doi: 10.1007/s10586-022-03776-z.

[8] M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," *Journal of Information Security and Applications*, vol. 50, Feb. 2020, doi: 10.1016/j.jisa.2019.102419.

[9] A. Arqane, O. Boutkhoum, H. Boukhriss, and A. E. Moutaouakkil, "Intrusion detection system using ensemble learning approaches: a systematic literature review," *International Journal of Online and Biomedical Engineering*, vol. 18, no. 13, pp. 160–175, Oct. 2022, doi: 10.3991/ijoe.v18i13.33519.

[10] Y. Wu, "DDos attack detection method based on machine learning," *Applied and Computational Engineering*, vol. 18, no. 1, pp. 88–95, Oct. 2023, doi: 10.54254/2755-2721/18/20230968.

[11] H. Zang, H. Kim, and J. Kim, "Blockchain-based decentralized storage design for data confidence over cloud-native edge infrastructure," *IEEE Access*, vol. 12, pp. 50083–50099, 2024, doi: 10.1109/ACCESS.2024.3383010.

[12] R. Lazzarini, H. Tianfield, and V. Charissis, "A stacking ensemble of deep learning models for IoT intrusion detection," *Knowledge-Based Systems*, vol. 279, Nov. 2023, doi: 10.1016/j.knosys.2023.110941.

[13] Z. Chen *et al.*, "Machine learning-enabled IoT security: open issues and challenges under advanced persistent threats," *ACM Computing Surveys*, vol. 55, no. 5, pp. 1–37, May 2023, doi: 10.1145/3530812.

[14] E. M. Maseno, Z. Wang, and H. Xing, "a systematic review on hybrid intrusion detection system," *Security and Communication Networks*, vol. 2022, pp. 1–23, May 2022, doi: 10.1155/2022/9663052.

[15] U. S. Musa, M. Chhabra, A. Ali, and M. Kaur, "Intrusion detection system using machine learning techniques: a review," in *2020 International Conference on Smart Electronics and Communication (ICOSEC)*, Sep. 2020, pp. 149–155, doi: 10.1109/ICOSEC49089.2020.9215333.

[16] O. Okporokpo, F. Olajide, N. Ajienka, and X. Ma, "Trust-based approaches towards enhancing IoT security: a systematic literature review," *Computer Science & Information Technology*, Nov. 2023, pp. 25–46, doi: 10.5121/csit.2023.132103.

[17] O. O. Oladimeji, A. B. Kayode, A. A. Olusola, and A. O. Isaiah, "Evaluation of selected stacked ensemble models for the optimal multi-class cyber-attacks detection," *International Journal on Cyber Situational Awareness*, vol. 5, no. 1, pp. 26–48, Jan. 2021, doi: 10.22619/IJCSA.2020.100132.

[18] D. S. Rajput and A. K. Upadhyay, "Enhanced network defense: optimized multi-layer ensemble for DDoS attack detection," *International Journal of Experimental Research and Review*, vol. 46, pp. 253–272, Dec. 2024, doi: 10.52756/ijerr.2024.v46.020.

[19] V. Safronov, I. Bostan, N. Allott, and A. Martin, "How memory-safe is IoT? assessing the impact of memory-protection solutions for securing wireless gateways," in *14th International Conference on the Internet of Things*, Nov. 2024, pp. 261–266, doi: 10.1145/3703790.3703820.

[20] M. M. Merlec and H. P. In, "Blockchain-based decentralized storage systems for sustainable data self-sovereignty: a comparative study," *Sustainability*, vol. 16, no. 17, Sep. 2024, doi: 10.3390/su16177671.

[21] R. Soleymanzadeh, M. Aljasim, M. W. Qadeer, and R. Kashef, "Cyberattack and fraud detection using ensemble stacking," *Artificial Intelligence*, vol. 3, no. 1, pp. 22–36, Jan. 2022, doi: 10.3390/ai3010002.

[22] W. F. Urmi *et al.*, "A stacked ensemble approach to detect cyber attacks based on feature selection techniques," *International Journal of Cognitive Computing in Engineering*, vol. 5, pp. 316–331, 2024, doi: 10.1016/j.ijcce.2024.07.005.

[23] T. Zoppi and A. Ceccarelli, "Prepare for trouble and make it double! supervised – unsupervised stacking for anomaly-based intrusion detection," *Journal of Network and Computer Applications*, vol. 189, Sep. 2021, doi: 10.1016/j.jnca.2021.103106.

[24] T. Sasi, A. H. Lashkari, R. Lu, P. Xiong, and S. Iqbal, "An efficient self attention-based 1D-CNN-LSTM network for IoT attack detection and identification using network traffic," *Journal of Information and Intelligence*, vol. 3, no. 5, pp. 375–400, Sep. 2025, doi: 10.1016/j.jiixd.2024.09.001.

[25] E. Alalwany, B. Alsharif, Y. Alotaibi, A. Alfahaid, I. Mahgoub, and M. Ilyas, "Stacking ensemble deep learning for real-time intrusion detection in IoMT environments," *Sensors*, vol. 25, no. 3, Jan. 2025, doi: 10.3390/s25030624.

[26] K. P. Sharma *et al.*, "Interpretable intrusion detection for IoT environments using a self-attention-based explainable AI framework," *Scientific Reports*, vol. 15, no. 1, Nov. 2025, doi: 10.1038/s41598-025-23750-0.

[27] P. Khordadpour and S. Ahmadi, "Security and privacy enhancing in blockchain-based IoT environments via anonym auditing," 2024, *arXiv:2403.01356*.

[28] A. Ba and M. Adda, "Intrusion detection in IIoT using machine learning," *Procedia Computer Science*, vol. 251, pp. 265–272, 2024, doi: 10.1016/j.procs.2024.11.109.

[29] E. M. Maseno and Z. Wang, "Intrusion detection in IoT using ensemble approach," in *5th International Symposium on Advanced Technologies and Applications in the Internet of Things (ATAIT 2023)*, 2023, pp. 15–24.

## BIOGRAPHIES OF AUTHORS

**Zobayer Alam** is an undergraduate student of Computer Science and Engineering in the Department of Computer Science, Faculty of Science and Technology at American International University–Bangladesh. His research interests include network security, IDSs, and IoT security. He is also a cybersecurity enthusiast. He can be contacted at email: zobayeralam1025@gmail.com.

**Arnab Bishakh Sarker** is an undergraduate student in Computer Science and Engineering at Department of Computer Science, Faculty of Science and Technology, American International University–Bangladesh. He is a full stack developer proficient in .NET framework, ReactJS, and NestJS, with interests in ML, computer vision, and pattern recognition. His research interests include intrusion detection in IoT environments, ensemble learning, and secure distributed systems. He can be contacted at email: bishakh99@outlook.com.

**Jariatun Islam** is a B.Sc. student of Computer Science and Engineering in the Department of Computer Science, Faculty of Science and Technology at American International University–Bangladesh. Her research interests include data science, ML, web development, and volunteering. She can be contacted at email: zariatunislam@gmail.com.

**Sifat Rahman Ahona** is an assistant professor in the Department of Computer Science, Faculty of Science and Technology at American International University–Bangladesh, where she completed both her B.Sc. in Computer Science and Engineering and M.Sc. in Computer Science. Her research interests include computer vision, data mining, artificial intelligence, deep learning, and ML. She has received summa cum laude at the 18th Convocation and the Dean's List Award for academic excellence. She can be contacted at email: ahona@aiub.edu.