Machine Learning Evaluation of the HiTar-2024 Dataset for Intrusion Detection in Smart Manufacturing Environments

Adeeb Alhomoud
College of Computing and Informatics, Saudi Electronic University, Riyadh 11673, Saudi Arabia

***Abstract*—This study presents an investigation of the HiTar-2024 dataset performed in terms of the distribution of label attack types and the distribution of attacks by protocol, normal, and Denial of Service (DoS) connections over time. The investigation carried out a performance evaluation of the HiTar-2024 dataset using a machine learning approach to classify benign and malicious activities, based on BayesNet, Logistic, IBk, Multiclass, PART, and J48 classifiers. It was found that the HiTar-2024 dataset can serve as a training set for an anomaly-based intrusion detection system (IDS) in a smart manufacturing environment to detect normal and malicious activities. Furthermore, an anomaly-based IDS using the HiTar-2024 dataset is able to group malicious activities into Probing, Remote-to-Local, User-to-Root, and DoS attacks.**

Keywords*—Supervised learning; intrusion detection system; security; smart manufacturing

I. Introduction

Today, intrusion detection systems (IDS) play a crucial role in detecting and stopping sophisticated attacks and malware propagation, especially in the Internet-of-Things (IoT) environment known as resource constrained devices. The smart grid is a particular case of an IoT environment in which any activity disruption leads to catastrophic damage to the fabrication process of smart manufacturing, market, service provider, and energy distribution systems [1]. The IDS taxonomy in the IoT environment contains four categories: signature-based [2], anomaly-based [3], specification-based [4], and hybrid detection mechanisms [5]. An anomaly-based IDS relies on machine learning (ML) and deep learning (DL) techniques. An optimized IDS ML based on a support vector machine (SVM) has been proposed by fitting the SVM kernel function and the complexity parameter, leading to a maximal value of the accuracy of the model [6]. A DL classifier called a Bidirectional long short-term memory (Bi-LSTM) model has achieved accuracy of up to 97.93% [7].

The taxonomy of attacks based on KDD'99 is composed of four classes: Probing, Remote-to-Local (R2L), User-to-Root (U2R), and Denial-of-Service (DoS) attacks [8]. In a Probing attack, the hacker will gather the maximum information possible about the target IoT victim vulnerability [9]. The hacker will then exploit this vulnerability to gain R2L access to the target IoT device or network gateway [10]. In addition, the aim of a U2R attack is, for instance, to gain root privileges through a memory overflow mechanism [11]. Finally, a hacker can stop or degrade the performance of a target IoT network through a DoS or Distributed DoS (DDoS) attack [12].

In Industrial IoT (IIoT) environments, the creation and thorough analysis of dedicated datasets is fundamental to developing an effective IDS. Standard network-security datasets, such as those developed for traditional IT networks, often fail to capture the unique characteristics of IIoT architectures, which include constrained devices, heterogeneous sensors, specialized communication protocols, and cyber-physical interactions. Recent research by Essop et al. [13] emphasizes that many existing IoT and IIoT datasets fail to capture the diversity of real-world conditions, as they often lack events that encompass multiple benign and attack scenarios. Furthermore, these datasets typically omit critical contextual information, such as sensor measurement data, network-level traffic characteristics, and device behavior within operational environments. The introduction of the Canadian Institute for Cybersecurity Advanced Persistent Threat for IIoT (CICAPT-IIoT) dataset addresses this gap by including network logs and provenance data across more than 20 advanced persistent threat techniques in an IIoT testbed [14]. Through dataset construction and detailed analysis of attack distributions, IoT device behavior, protocol usage distribution, and temporal patterns, researchers can better understand threat vectors related to the IIoT environment. This process could enable the design of IDS models based on ML that are both accurate and practical in various applications, such as smart manufacturing, energy distribution, and critical infrastructure settings. Without the existence of those targeted datasets, the IDS models generated will suffer from poor generalizability, class imbalance, or misalignment with real-world IIoT networks. The main contributions of this study are summarized as follows:

- An investigation of the HiTar-2024 dataset generated from a smart manufacturing simulation environment.
- A HiTar-2024 dataset performance evaluation based on BayesNet, Logistic, Instance-Based k-Nearest Neighbors (IBk), Multiclass, partial decision trees (PART), and J48 classifiers.
- The HiTar-2024 dataset can serve as a training set to detect Probing, R2L, U2R, and DoS attacks before its deployment online in a smart manufacturing environment.

The rest of the study is composed into four sections: Section II presents the state-of-the-art of the existing IIoT dataset in the literature. Section III describes the HiTar-2024 dataset by considering its most important features. Section IV investigates a performance evaluation of the HiTar-2024 dataset and presents a confusion matrix of the classifiers used. Section V presents the results and discussion. Section VI draws the conclusion.

II. State-of-the-Art

Table I compares IIoT datasets in the literature as a function of year creation, dataset name, number of features, attack classes, and a brief description.

TABLE I. IIoT Datasets in the Literature

Year	Dataset Name	Number of Features	Attack Classes	Description
2020	TON_IoT [15]	127	Normal, DDoS, DoS, Injection, MiM, Password, XSS, Scanning.	A telemetry dataset is proposed for both IoT and IIoT applications.
2021	WUSTL-IIOT-2021 [16]	41	Command Injection, DoS, Reconnaissance, Backdoor.	WWUSTL-IIoT-2021 consists of network data of Industrial Internet of Things to be used in cybersecurity research.
2022	Edge IIoTset [17]	61	DoS/DDoS attacks, Information gathering, MiM attacks, Injection attacks, Malware attacks.	The Edge-IIoTset dataset generated could be used by an IDS based on centralized and federated learning.
2024	CICAPT-IIoT [18]	140	Collection, Exfiltration, Command and Control, Persistence, Discovery, Credential Access, Lateral Movement and Defence Evasion.	An APT attack dataset captured within the IIoT environment.
2024	HiTar-2024 [19]	39	Probing, R2L, U2R and DoS	HiTar-2024 is a dataset generated from the AREZZO environment that represents a simulated industrial manufacturing process
2025	DataSense: CIC IIoT [20]	94	Benign, DoS, DDoS, MiTM, Mirai, Bruteforce	A realistic IIoT testbed dataset with synchronized sensor and network data for anomaly detection.

Important IIoT datasets that are currently available in the literature are compared in this work along with the main distinctions between them, such as the number of features extracted, the variety of attacks that are possible, and their suitability for cybersecurity research in the smart manufacturing field. A summary of the early attempts to create comprehensive datasets is provided. One of the first comprehensive datasets was the TON_IoT [15], which had 127 features and covered a wide range of attack types, such as DDoS, DoS, injection, and password attacks, making it suitable for anomaly detection based on telemetry for IoT/IIoT networks.

The 41 features of WUSTL-IIoT-2021 [16], on the other hand, cover injection, DoS, reconnaissance, and backdoor attacks, all of which are essential for assessing targeted IIoT attacks. The Edge IIoTset dataset contains 61 new features and could be used by an IDS based on centralized and federated learning [17]. The CICAPT-IIoT dataset was introduced in 2024 [18]. The latter dataset contains 20 kinds of attacks that can be carried out in an IIoT environment, including collection, exfiltration, command and control, persistence, discovery, credential access, lateral movement, and defence evasion classes. Another, HiTar-2024, is a dataset generated from AREZZO environment LOG files that represents a simulated industrial manufacturing process in an IIoT context. The HiTar-2024 dataset includes 39 features, 15,842 instances, and four attack classes: Probing, R2L, U2R, and DoS attacks [19]. Furthermore, the HiTar-2024 was created following a rigorous approach involving data preprocessing, feature extraction, the class imbalance with SMOTE, and a test option that was employed for model robustness. Finally, the DataSense CIC-IIoT Dataset 2025, also developed by the CIC, represents one of the most recent and comprehensive datasets designed for intrusion detection in IIoT environments. Built from a realistic IIoT testbed, it integrates both sensor telemetry and network traffic data, providing a synchronized view of industrial device behavior and communication flows [20].

III. Description of the Dataset

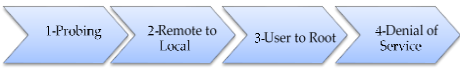


Fig. 1. Attack class sequence order

Fig.1 represents the chronological attack class order used in HiTar-2024 generation. The process started by collecting information about the target victim (probing class). In the attack scenario, NMAP commands were used to perform a probing attack. Subsequently, a remote connection to the target AREZZO system was attempted (R2L class). In addition, Rlogin, Secure Shell (SSH), and Telnet commands were examined as the R2L attack class. A U2R class is a tentative attempt to gain Root permission. Sudo, Su, and FTP commands were then investigated to generate U2R attacks. Finally, a DoS attack was launched in order to stop or degrade a service or network activity related to the AREZZO environment (DoS class). The DoS attack was investigated through the Hping3 command.

The following section presents the distribution of label attack types and the distribution of attacks by protocol, normal, and DoS connections over time in the HiTar2024 dataset using the JupyterLab environment. The pie chart in Fig. 2 shows the distribution of label attack types in the HiTar2024 dataset. Each

slice represents how frequently a specific behavior or attack type occurs in the dataset. As shown above, Probing attacks represent 65.6% of the total attacks, Normal behavior represents 25.6% of all traffic, R2L attacks are 4.9% of the total traffic, U2R attacks are 3.1% of the total attacks, and DoS attacks represent 0.8% of the global collected traffic. This indicates that the majority of the records in the dataset involve scanning or probing activities, whereby an attacker tries to gather information about the system (e.g., port scanning or network mapping). The dataset is highly imbalanced, with one class (Probing_Attack) dominating. As a consequence, a more specific technique (e.g., class weighting, oversampling rare attacks such as a DoS, or undersampling probing attacks) is necessary to avoid biased detection models.

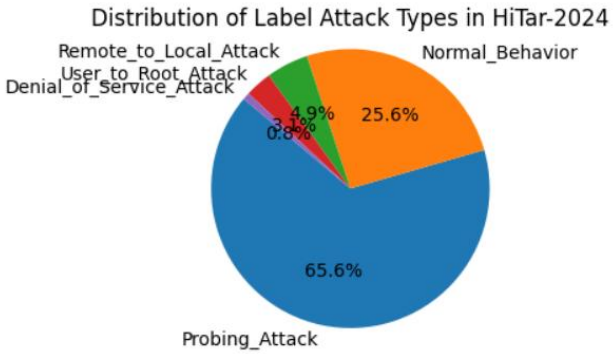


Fig. 2. Distribution of label attack types in the HiTar-2024 dataset.

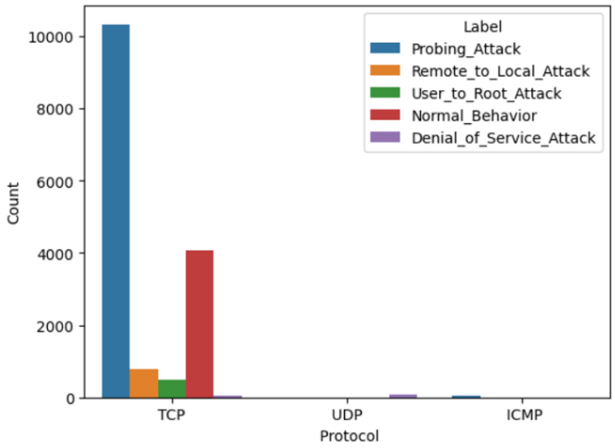


Fig. 3. Distribution of attacks by protocol in the HiTar-2024 dataset.

Fig. 3 shows the distribution of different attack categories and normal traffic across network protocols (Transmission Control Protocol [TCP], User Datagram Protocol [UDP], and Internet Control Message Protocol [ICMP]) within the HiTar-2024 dataset. The results show a strong imbalanced usage of protocol types, indicating that the majority of both benign and malicious activities are transmitted over TCP. Specifically, Probing attacks are the most frequent, exceeding 10,000 TCP instances, suggesting extensive scanning or probing activity targeting IoT devices. R2L and U2R attacks also occur mainly over TCP, although at considerably lower frequencies, which aligns with these attacks typically requiring a stable connection to be established in order to exploit system vulnerabilities and weaknesses. Finally, Normal traffic is mainly transmitted over TCP, with 4000 instances, reinforcing the central role of TCP in typical IIoT communication patterns. To draw a conclusion, the imbalance across protocols indicates that IIoT intrusion detection models should incorporate knowledge of protocol features, particularly identifying TCP traffic patterns.

Fig. 4 shows the Normal and DoS connections over time in the HiTar-2024 dataset. The pic connections represent the DoS attacks when compared to Normal activity (without pic connections).

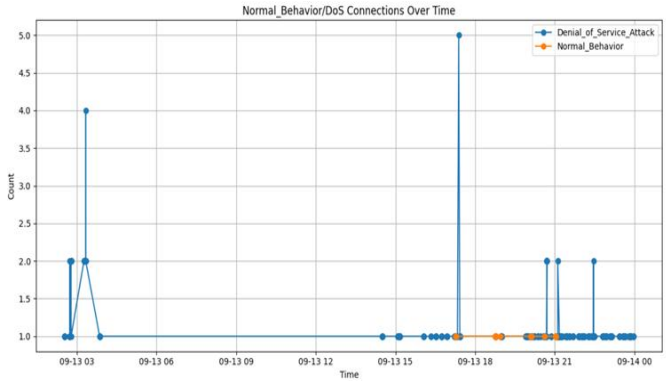


Fig. 4. Normal and DoS connections over time in the HiTar-2024 dataset.

As seen in Fig. 4, the various spikes represent DoS attacks, and the line below represents Normal activity. Fig. 4 indicates a high-intensity attack phase, which is likely to correspond to a targeted flooding attempt. In contrast, Normal behavior appears only sporadically during this interval. Between 06:00 and 15:00 on 09-13, the traffic stabilizes, and DoS activity is minimal, suggesting either a paused attack or normal operational conditions in the IIoT network. However, around 18:00, another sharp spike in DoS activity is observed, indicating a second attack wave, again with significantly higher frequency than the normal traffic. Approaching the later period (around 20:00 to 23:00), both Normal and DoS connections appear intermixed, but DoS packets still dominate in terms of count. This behavior may indicate either background scanning/flooding attempts or a sustained low-rate persistent DoS attack intended to evade detection.

IV. Dataset Performance Evaluation

A structured comparison of the different classifiers tested for the HiTar-2024 dataset is shown in Table II. BayesNet, Logistic regression, IBk, Multiclass classification techniques, PART, and J48 (also known as the C4.5 algorithm) are included in the table. The computational complexity, main benefits in the classification context, and potential drawbacks that could affect the suitability of each classifier for IIoT threat detection are described in the table.

TABLE II. CLASSIFIERS USED IN HITAR-2024 PERFORMANCE EVALUATION

Classifier Name	Description	Complexity	Advantages	Disadvantages
BayesNet [19,22,24]	A probabilistic graphical model that uses Bayesian networks for classification.	Low	- Efficient for small datasets - Handles missing data well - Interpretable results	- Assumes conditional independence - Less effective with complex, high-dimensional data
Logistic [19,21,23,24]	A statistical model that predicts categorical outcomes using a logistic function.	Low to Medium	- Works well with linear decision boundaries - Robust against overfitting when regularization is applied	- Struggles with non-linearly separable data - Performance drops with highly imbalanced datasets
IBk [19,24]	A lazy learning algorithm that classifies based on the majority class of nearest neighbors.	High	- Simple and intuitive - Effective for non-linear relationships - No training phase, adapts dynamically	- Computationally expensive for large datasets - Sensitive to noise and irrelevant features
Multiclass [19,24]	A classification method that extends binary classifiers to handle multiple classes simultaneously.	Medium to High	- Efficient for multi-label problems - Can be integrated with various base classifiers	- Performance depends on the decomposition method chosen - Computational cost increases with class complexity
PART [19,21,24]	A rule-based classifier that builds decision rules from partial decision trees.	Medium	- Generates interpretable rules - Less prone to overfitting compared to full decision trees	- Can be computationally expensive - Rule generation may not always be optimal
J48 [19,21,24]	A decision tree-based classifier that recursively splits data based on attribute values.	Medium to High	- Easily interpretable - Handles both numerical and categorical data - Can handle missing values	- Prone to overfitting - Can become complex and hard to optimize

The following classifier configurations were used to perform system performance evaluation:

weka.classifiers.bayes.BayesNet -D -Q weka.classifiers.bayes.net.search.local.K2 -- -P 1 -S BAYES -E weka.classifiers.bayes.net.estimate.SimpleEstimator -- -A 0.

weka.classifiers.functions.Logistic ---R 1.0E-8 -M -1 -num-decimal-places 4

weka.classifiers.lazy.IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""

weka.classifiers.rules.PART -C 0.25 -M 2

weka.classifiers.meta.MultiClassClassifier -M 0 -R 2.0 -S 1 -W

weka.classifiers.trees.J48 -C 0.25 -M 2

The methodology combined three effective techniques to address overfitting in the tabular data in the HiTar-2024 dataset, including balancing the minority class using SMOTE (Synthetic Minority Over-sampling Technique), feature selection by selecting an attribute evaluator, and a search method and k-fold cross-validation in test option processing. Regarding feature selection, the CfsSubsetEval and BestFirst were selected as the attribute evaluator and search method, respectively. The choice of the CfsSub-setEval feature selection method was motivated by this suitability for tabular data, it produces a compact, nonredundant subset, and improves model generalization.

Furthermore, imbalanced classes are handled by SMOTE, such as the U2R and DoS attack classes. The following setting is used in the preprocessing supervised filter based on the instance:

weka.filters.supervised.instance SMOTE -C 0 -K 5 -P 400.0 -S 1

Since the minority class DoS has a number of instances of 130, the percentage parameter of SMOTE is selected to 400% to reach the number of U2R instances equal to 491.

A k-fold cross-validation test option was also applied, partitioning the data into k subsets (or folds), training the model on k-1 subsets, and testing it on the held-out subset. The simulation was carried out for k equal to 10 for the cross-validation test option.

weka.attributeSelection.CfsSubsetEval -P 1 -E 1

CfsSubsetEva evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them.

weka.gui.attributeSelection.BestFirst -D 1 -N 5

BestFirst searches the space of attribute subsets by using greedy hill climbing augmented with a backtracking facility.

A comprehensive set of experiments with common ML models using BayesNet, Logistic, IBk, Multiclass, PART, and J48 with the HiTar-2024 dataset demonstrates high accuracy, precision, recall, and F1-scores, as well as around 99% accuracy [19]. For example, the authors in [19] made a comparison between Simple Logistic, PART, and J48 classifiers using the NSL-KDD dataset. They also examined the detection of different attack types involving DoS, R2L, U2R, and Probing attacks and considered time when building the model. The authors in [22] tested the BayesNet classifier using NSL-KDD and found higher true positive (TP) and lower false positive (FP) rates when compared to Naïve Bayes and updateable versions. Other authors proposed Logistic regression, decision trees, k-Nearest Neighbors, and Naïve Bayes classifiers on the WUSTL-IIOT-2021 dataset [23]. The latter study gives a context for how IBk and Logistic classifiers perform under IIoT network traffic data conditions.

Table III presents a performance comparison of classifiers for HiTar-2024 as a function of accuracy, Kappa statistic, mean absolute error (MAE), root mean square error (RMSE), and relative absolute error (RAE) metrics.

As shown in Table III, all classifiers demonstrate remarkably high accuracy, exceeding 98%, which indicates that the proposed dataset and feature selection approach enable highly reliable classification. Among the models evaluated, the IBk algorithm achieved the highest accuracy, reaching 99.993%, accompanied by a Kappa statistic of 0.999, MAE of 0.0001, and RMSE of 0.005. These results highlight the robustness and consistency of the IBk model, confirming its superior ability to minimize both MAE and RMSE errors. The Logistic regression model also achieved competitive results, with an accuracy of 99.974%, a Kappa of 0.9995, MAE of 0.0001, and RMSE of 0.0101. The J48 decision tree and Multiclass classifiers followed closely, each exceeding 99.9% accuracy, and demonstrating excellent agreement levels with a Kappa of around 0.999.

TABLE III. CLASSIFIER PERFORMANCE COMPARISON FOR HITAR-2024

Classifier	Accuracy (%)	Kappa Statistic	Mean Absolute Error	Root Mean Square Error	Relative Absolute Error
BayesNet	98.668	0.973	0.005	0.069	2.471
Logistic	99.974	0.9995	0.0001	0.0101	0.053
IBk	99.993	0.999	0.0001	0.005	0.0623
PART	99.899	0.998	0.019	0.2903	6.2892
Multiclass	99.949	0.999	0.0004	0.0143	0.2094
J48	99.968	0.999	0.0002	0.011	0.123

In contrast, although the PART and BayesNet classifiers also showed good results with an accuracy of 99.899% and 98.668%, respectively, they showed higher error values regarding an RMSE equal to 0.2903 and 0.069 and an RAE of 6.2892 and 2.471, suggesting a slightly lower prediction precision compared to the top-performing IBk model. Nonetheless, their Kappa statistics of above 0.97 confirm consistent and reliable performance over classes. In conclusion, the IBk and Logistic regression classifiers can be identified as the most suitable models for the HiTar-2024 dataset, achieving excellent classification with minimal prediction error. These findings confirm that the adopted approach is effective at enhancing classification accuracy and maintaining high robustness and generalization capability.

Table IV presents a performance comparison between recent IIoT datasets and HiTar-2024 in terms of task, typical accuracy, precision, recall, and F1-score.

TABLE IV. CLASSIFIER PERFORMANCE COMPARISON BETWEEN HITAR-2024 AND OTHER IIOT DATASETS

Classifier	Task	Typical accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CICAPT-IIoT	APT-style detection	~93.8	~93	~95	~94
Edge-IIoTset	Multiclass attack detection	95–99.97	~95–99.3	~95–99.3	~95–99.3
DataSense (CIC IIoT)	Binary & multiclass IIoT	~98–99	~98	~98	~98
HiTar-2024	Multiclass attack detection	98.668–99.993	99.1–99.9	99.0– 99.9	99.0–99.9

As shown in Table IV, HiTar-2024 outperforms other recent IIoT datasets (i.e., CI-CAPT-IIoT, Edge-IIoTset, and DataSense (CIC IIoT)) in terms of typical accuracy, precision, recall, and F1-score, motiving its suitability for smart manufacturing.

Next, the performance evaluation was conducted using a confusion matrix for the various classifiers studied in a WEKA simulation environment using the same setting as in [17].

Fig. 5(a) presents a colored confusion matrix heatmap for the BayesNet classifier, in which the predicted values are shown in the columns and the defined values of the dataset in the rows. The confusion matrix for the BayesNet classifier presents the following insights:

- Probing attack: The classifier correctly predicted 10,261 instances of Probing attacks, with a maximum TP value (i.e., equal to 100%) and a minimum FP value (i.e., equal to 0%).
- R2L attack: The classifier correctly predicts 754 instances of R2L attacks (representing a TP rate of 87.98%). In addition, 102 instances were predicted as probing attacks and one instance predicted as a U2R attack (representing an FP rate of 12.02%).
- U2R attack: The classifier correctly predicted 481 instances of U2L attacks (representing a TP rate of 94.49%), but 16 instances were predicted as probing attacks and 12 instances as R2L attacks (representing an FP rate of 5.51%).
- Normal behavior: The classifier correctly predicted 4058 instances of normal behavior (representing a TP rate of 99.68%). However, 7 instances were predicted as Probing attacks and 6 as R2L attacks (representing an FP rate of 0.32%).
- DoS attack: The classifier successfully predicted 130 instances of DoS attacks (representing a TP rate of 90.27%). Two instances were predicted as Probing

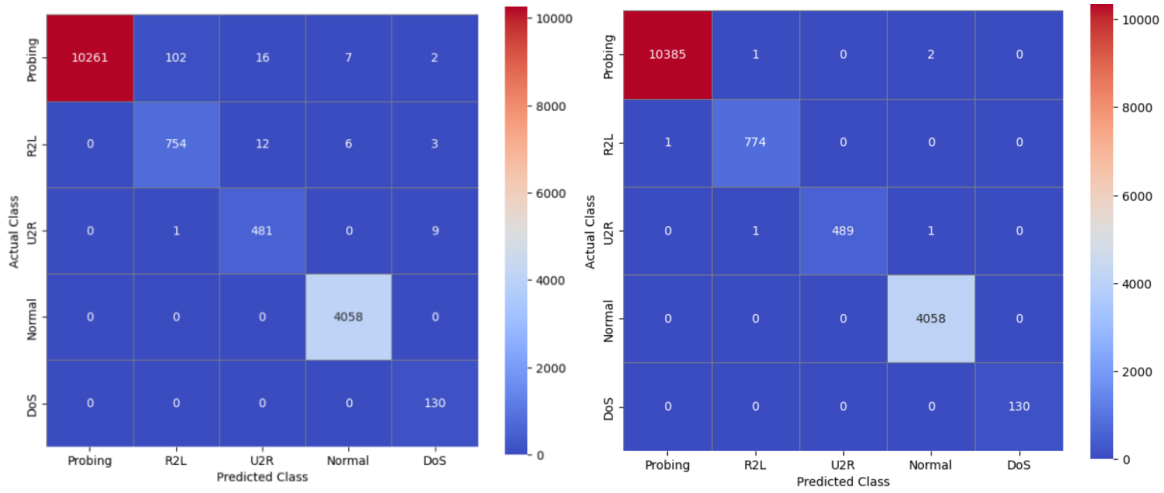
attacks, 3 instances were predicted as R2L attacks, and 9 instances were predicted as U2R attacks (representing an FP rate of 9.73%).

In summary, the BayesNet classifier seems to excel in predicting Probing attacks and Normal behavior, but performs less well in predicting R2L, U2R, and DoS attacks.

Fig. 5(b) presents a colored confusion matrix heatmap for the Logistic classifier, in which the predicted values are shown in the columns and the defined values of the dataset in the rows. The confusion matrix for the Logistic classifier enables the following insights:

- Probing attack: The classifier correctly predicted 10,385 instances of probing attacks (with a TP value equal to 99.99%) and 1 instance was predicted as an R2L attack (with an FP equal to 0.01%).
- R2L attack: The classifier correctly predicted 774 instances of R2L attacks (representing a TP rate of 99.74%). One instance was predicted as a Probing attack and 1 instance as a U2R attack (representing an FP rate of 0.26%).
- U2R attack: The classifier correctly predicted 489 instances of U2R attacks, with a maximum TP value (i.e., equal to 100%) and a minimum FP value (i.e., equal to 0%).
- Normal behavior: The classifier correctly predicted 4058 instances of Normal behavior (representing a TP rate of 99.92%). However, 2 instances were predicted as Probing attacks and 1 instance was predicted as a U2R attack (representing an FP rate of 0.08%).
- DoS attack: The classifier correctly predicted 130 instances of DoS attacks, with a maximum TP value (i.e., equal to 100%) and a minimum FP value (i.e., equal to 0%).

In conclusion, the Logistic classifier seems to excel in predicting U2R and DoS attacks, but performs less well in predicting Probing and R2L attacks and Normal behavior.



(a) (b)

Fig. 5. (a) Colored confusion matrix heatmap for the BayesNet classifier; (b) Colored confusion matrix heatmap for the Logistic classifier.

Fig. 6(a) presents a colored confusion matrix heatmap for the IBk classifier, in which the predicted values are shown in the columns and the defined values of the dataset are shown in the rows. The confusion matrix for the IBk classifier enables the following insights:

- Probing attack: The classifier correctly predicted 10,388 instances of probing attacks, with a maximum TP value (i.e., equal to 100%) and a minimum FP value (i.e., equal to 0%).
- R2L attack: The classifier correctly predicted 775 instances of R2L attacks, with a maximum TP value (i.e., equal to 100%) and a minimum FP value (i.e., equal to 0%).
- U2R attack: The classifier correctly predicted 491 instances of U2R attacks (with a TP rate equal to 99.79%) and 1 instance was predicted as a DoS attack (with an FP equal to 0.03%).
- Normal behavior: The classifier correctly predicted 4058 instances of Normal activity, with a maximum TP value (i.e., equal to 100%) and a minimum FP value (i.e., equal to 0%).
- DoS attack: The classifier correctly predicted 129 instances of U2R attacks, with a maximum TP value (i.e., equal to 100%) and a minimum FP value (i.e., equal to 0%).

In summary, the IBk classifier seems to excel in predicting Probing, R2L, and DoS attacks and Normal behavior, but performs slightly less well in predicting U2R attacks.

Fig. 6(b) presents the confusion matrix for the Multiclass classifier, in which the predicted values are shown in the columns and the defined values of the dataset in the rows.

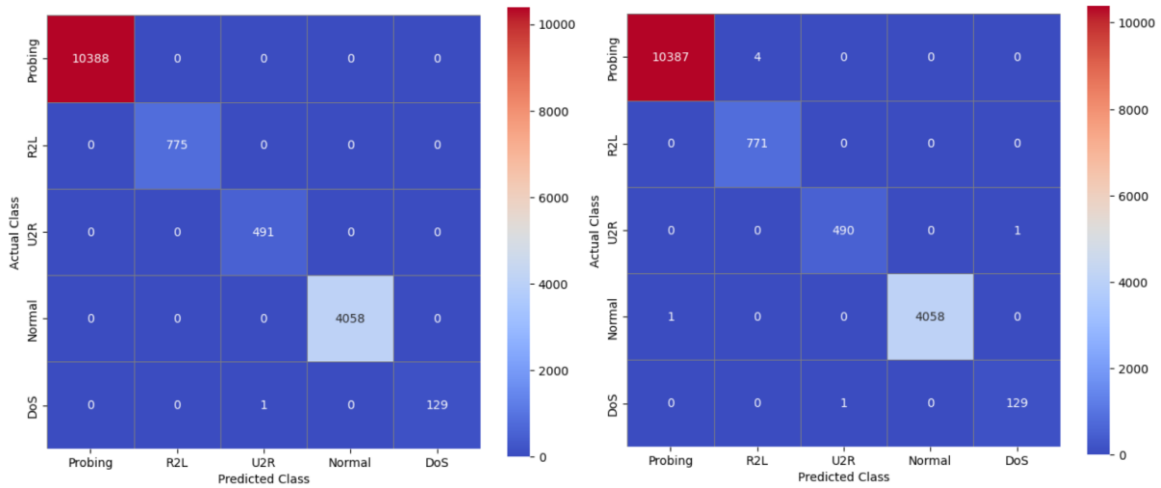


Fig. 6. (a) Colored confusion matrix heatmap for the IBk classifier; (b) Colored confusion matrix heatmap for the Multiclass classifier.

The confusion matrix for the Multiclass classifier presents the following insights:

- Probing attack: The classifier correctly predicted 10,387 instances of Probing attacks (with a TP rate equal to 99.96%) and 4 instances predicted as R2L attacks (with an FP rate equal to 0.04%).
- R2L attack: The classifier correctly predicted 771 instances of R2L attacks, with a maximum TP value (i.e., equal to 100%) and a minimum FP value (i.e., equal to 0%).
- U2R attack: The classifier correctly predicted 490 instances of U2R attacks (with a TP rate equal to 99.79%) and 1 instance predicted as a DoS attack (with an FP rate equal to 0.21%).
- Normal behavior: The classifier correctly predicted 4058 instances of U2R attacks (with a TP rate equal to 99.97%) and 1 instance predicted as a Probing attack (with an FP rate equal to 0.03%).
- DoS attack: The classifier correctly predicted 129 instances of U2R attacks (with a TP rate equal to 99.23%) and 1 instance predicted as a U2R attack (with an FP rate equal to 0.73%).

In summary, the Multiclass classifier seems to excel in predicting R2L attacks. However, it performs slightly less well in predicting Probing, U2R, and DoS attacks.

Fig. 7(a) presents a colored confusion matrix heatmap for the PART classifier, in which the predicted values are shown in the columns and the defined values of the dataset in the rows.

The confusion matrix for the PART classifier presents the following insights:

- Probing attack: The classifier correctly predicted 10,387 instances of Probing attacks (with a TP rate equal to 99.96%) and 1 instance predicted as a DoS attack (with an FP rate equal to 0.04%).
- R2L attack: The classifier correctly predicted 772 instances of R2L attacks (with a TP rate equal to 98.84%) and 9 instances predicted as a Probing attack (with an FP rate equal to 1.16%).
- U2R attack: The classifier correctly predicted 491 instances of DoS attacks, with a maximum TP value (i.e., equal to 100%) and a minimum FP value (i.e., equal to 0%).
- Normal behavior: The classifier correctly predicted 4058 instances of U2R attacks (with a TP rate equal to 99.92%) and 3 instances predicted as R2L attacks (with an FP rate equal to 0.08%).
- DoS attack: The classifier correctly predicted 129 instances of U2R attacks (with a TP rate equal to 98.47%) and 2 instances predicted as Probing attacks (with an FP rate equal to 1.53%).

In summary, based on the confusion matrix above, the PART classifier seems to excel in predicting U2R attacks. However, the number of false alarms is equal to 1, 2, 3, and 9, respectively, for Probing, DoS, Normal, and R2L activities.

Fig. 7(b) shows a colored confusion matrix heatmap for the J48 classifier, in which the predicted values are shown in the columns and the defined values of the dataset in the rows.

The confusion matrix for the J48 classifier presents the following insights:

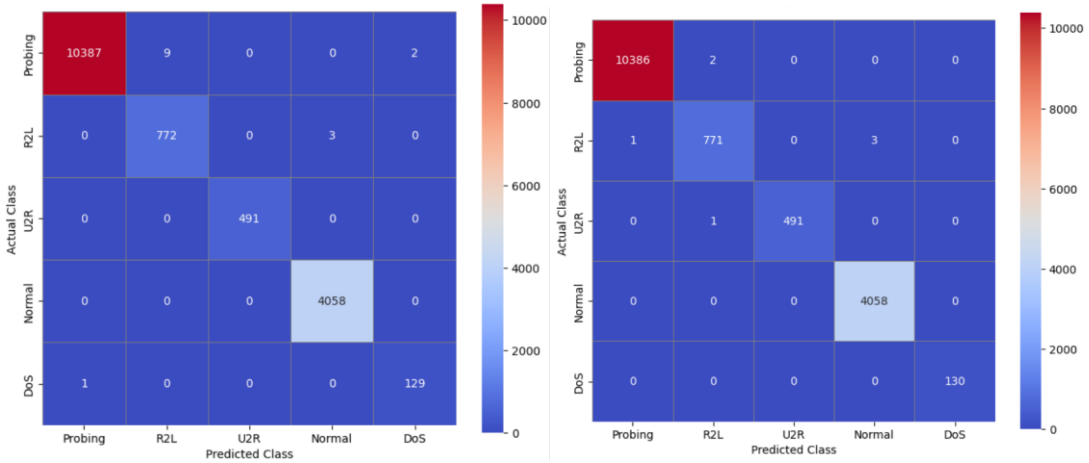
- Probing attack: The classifier correctly predicted 10,386 instances of Probing attacks (with a TP rate equal to 99.99%) and 1 instance predicted as an R2L attack (with an FP rate equal to 0.01%).
- R2L attack: The classifier correctly predicted 771 instances of R2L attacks (with a TP rate equal to 99.74%) and 2 instances predicted as Probing attacks (with an FP rate equal to 0.26%).
- U2R attack: The classifier correctly predicted 491 instances of U2R attacks, with a maximum TP value (i.e.,

equal to 100%) and a minimum FP value (i.e., equal to 0%).

- Normal behavior: The classifier correctly predicted 4058 instances of U2R attacks (with a TP rate equal to 99.92%) and 3 instances predicted as R2L attacks (with an FP rate equal to 0.08%).
- DoS attack: The classifier correctly predicted 130 instances of DoS attacks, with a maximum TP value (i.e., equal to 100%) and a minimum FP value (i.e., equal to 0%).

In summary, the J48 classifier seems to excel in predicting U2R and DoS attacks. However, the number of false alarms is equal to 1, 2, and 3, respectively, for the Probing, R2L, and Normal classes.

Table V compares the various classifiers studied in terms of Probing TP rate, R2L TP rate, U2R TP rate, Normal TP rate, DoS TP rate, average TP rate, and average FP rate.



(a) (b)

Fig. 7. (a) Colored confusion matrix heatmap for the PART classifier; (b) Colored confusion matrix heatmap for the J48 classifier.

TABLE V. TP AND FP RATES PER CLASSIFIERS STUDIED

Classifier	Probing TP Rate	R2L TP Rate	U2R TP Rate	Normal TP Rate	DoS TP Rate	Average TP Rate	Average FP Rate
BayesNet	100	87.98	94.49	99.68	90.27	94.48	5.52
PART	99.96	98.84	100	99.92	98.47	99.43	0.57
Multiclass	99.96	100	99.79	99.97	99.23	99.79	0.21
J48	99.99	99.74	100	99.92	100	99.93	0.07
Logistic	99.99	99.74	100	99.92	100	99.93	0.07
IBk	100	100	99.79	100	100	99.95	0.05

In Table V, the classifier performance is sorted as a function of the average TP rate and FP rate metrics as follows: IBk is the most effective classifier, followed by Logistic regression, then J48 followed by the Multiclass classifier, and the PART classifier is followed by BayesNet. Also as shown in Table V, the best-performing classifier is IBk and the one that performs least well is BayesNet in terms of average TP and FP rates. The FP rate is equal to 0.05 and 5.52 for IBk and BayesNet, respectively. The BayesNet classifier performance is given an overall TP rate of 94.48% and an FP rate of 5.52%.

V. RESULTS AND DISCUSSION

The comparative performance results in Table V above demonstrate clear differences in detection capability and false alarm behavior among the classifiers evaluated in the IIoT intrusion detection context. The BayesNet classifier achieved a reasonable average TP rate of 94.48%, but its relatively high FP rate (5.52%) indicates reduced reliability in high security environments, in which false alerts are costly. The PART classifier showed greater detection performance with an average TP rate of 99.43% and a lower FP rate (0.57%), benefiting from rule-based interpretability, although it remains slightly less accurate than tree-based models. The Multiclass classifier attained a more balanced detection level (99.79% average TP rate) with a low FP rate (0.21%), demonstrating its effectiveness across most attack categories. Among the decision-based methods, both J48 and Logistic regression achieved consistently strong performance (average TP = 99.93%, FP = 0.07%), making them highly suitable for deployment in resource-constrained IIoT edge devices due to their low computational overhead and model transparency. The IBk classifier obtained the highest overall detection accuracy (99.95% average TP rate) and the lowest FP rate (0.05%); however, its high computational cost and memory requirements during inference make it less appropriate for real-time or low-power IIoT nodes, but well-suited for centralized or offline security analytics. Overall, J48 and Logistic emerge as the most practical classifiers for real-time IIoT intrusion detection, whereas IBk is preferable when computational resources are not a constraint. In order to compare

this work with similar research in the literature, the TP rate is shown to be around 95% and the FP rate 4.87% when applying a BayesNet classifier to the NSL-KDD dataset [22]. In resource-constrained IIoT environments, the computational feasibility of ML classifiers depends primarily on inference complexity, memory footprint, and energy consumption. The IBk classifier is generally unsuitable for deployment on IIoT edge devices because it stores the entire training dataset and requires distance computations with all instances during inference, resulting in high memory usage and latency. In contrast, J48 exhibits strong feasibility due to its compact tree representation and low inference complexity, as classification requires only a simple tree traversal. Similarly, PART, a rule-based classifier derived from partial decision trees, offers efficient inference provided that the generated rule set remains limited in size. Multiclass Logistic regression is also viable for edge deployment when trained offline, since inference involves only linear computations over feature vectors, although its cost increases with feature dimensionality and number of classes. BayesNet, although probabilistically expressive, may impose higher memory and computational demands depending on network structure complexity and conditional probability tables. Overall, decision tree-based methods, such as J48, followed by PART and Logistic regression with feature selection, represent the most computationally viable solutions for real-time classification on constrained IIoT edge devices.

Finally, three concrete techniques were implemented to address dataset representativeness and evaluation realism, in terms of data collection (i.e., by oversampling rare attack classes such as the DoS class in HiTar-2024), dataset design (i.e., the HiTar-2024 dataset preserves timestamps in the first attribute to enable realistic splits and analysis), and through better metrics and reporting, such as using a confusion matrix and TP and FP rate per studied class. Although all six classifiers exceed 98% accuracy after SMOTE, the augmentation strategy appears to target only the DoS class, while the rare R2L and U2R classes remain insufficiently represented. Consequently, the reported accuracy may overestimate real-world robustness, and additional cross-dataset or out-of-distribution evaluation is needed to confirm generalization beyond the HiTar-2024 training partition.

VI. Conclusion

In this study, the HiTar-2024 dataset was evaluated using accuracy, Kappa statistic, MAE, RMSE, RAE, and confusion matrix-based performance metrics across several WEKA classifiers, namely BayesNet, Logistic, IBk, Multiclass, PART, and J48. The results indicate that the BayesNet classifier achieved an average TP rate of 94.48% along with an average FP rate of around 5.52%. The PART classifier achieved an average TP rate of 99.43% along with an average FP rate of 0.57%. The Multiclass classifier demonstrated an average TP rate of 99.79% with an average FP rate of 0.21%. An average TP rate of 99.93% along with an average FP rate of 0.07% was shown by the J48 and Logistic classifiers. The IBk classifier recorded the highest value in terms of an average TP rate of 99.95%, along with the lowest value for an average FP rate of 0.05%. These findings demonstrate that the IBk, followed by J48 and Logistic regression, constitute the most suitable set of classifiers for deployment in resource-constrained IIoT environments, offering an effective balance between classification performance and computational efficiency. These models achieve competitive detection accuracy while maintaining low memory consumption and fast inference time, making them particularly well adapted for real-time edge-level implementation.

The presence of high-intensity attack intervals followed by quiet periods in the traffic patterns suggests that the attacker employed intermittent flooding strategies, which is typical behavior in DoS scenarios targeting IIoT networks. Such bursts are designed to disrupt device availability while reducing the likelihood of long-term detection. In addition, the imbalance observed across communication protocols indicates that intrusion detection mechanisms in IIoT environments should incorporate protocol-aware learning features in future work, with particular significance given to TCP traffic behavior being considered as involving the absolute majority of malicious interactions. A potential perspective of this study could be the experiments, which validate detectors in real IIoT environments, the inclusion of advanced attacks such as APTs in the HiTar-2024 dataset, and the use of benchmark hybrid models that combine feature engineered ML, such as CNN/LSTM and CNN/GRU, to strengthen a detection intrusion system in an IIoT environment.

References

[1] Alomari MA, Al-Andoli MN, Ghaleb M, Thabit R, Alkawsi G, Alsayaydeh JAJ, Gaid ASA. Security of Smart Grid: Cybersecurity Issues, Potential Cyberattacks, Major Incidents, and Future Directions. Energies. 2025; 18(1):141. https://doi.org/10.3390/en18010141

[2] Alsoufi, M.A.; Razak, S.; Siraj, M.M.; Nafea, I.; Ghaleb, F.A.; Saeed, F.; Nasser, M. Anomaly-Based Intrusion Detection Systems in IoT Using Deep Learning: A Systematic Literature Review. Appl. Sci. 2021, 11, 8383.

[3] Anzila Saba, Amjad Rehman, Tariq Sadad, Hoshang Kolivand, Saeed Ali Bahaj, Anomaly-based intrusion detection system for IoT networks through deep learning model, Computers and Electrical Engineering, , 2022, Volume 99, 107810, ISSN 0045-7906, https://doi.org/10.1016/j.compeleceng.2022.107810.

[4] S. V. N. Santhosh Kumar and M. Selvi, A. Kannan, A Comprehensive Survey on Machine Learning-Based Intrusion Detection Systems for Secure Communication in Internet of Things, Computational Intelligence and Neuroscience, 2023.

[5] Yazan Otoum, Amiya Nayak, ASIDS: Anomaly and Signature Based IDS for the Internet of Things, Journal of Network and Systems Management, 2021 vol. 29(3),.

[6] A. Alhomoud, An Optimized Network Intrusion Detection System for Attack Detection based on Supervised Machine Learning Models in an Internet-of-Things Environment, International Journal of Advances in Soft Computing & Its Applications, 2023 Vol 15, Issue 2,

[7] Oueslati, N.E., Mrabet, H., Jemai, A. (2025). Intrusion Detection Using an Enhancement Bi-LSTM Recurrent Neural Network Model. In: Ben Hedia, B., Ghazel, M., Monsuez, B. (eds) Verification and Evaluation of Computer and Communication Systems. VECoS 2024. Lecture Notes in Computer Science, vol 15466. Springer, Cham. https://doi.org/10.1007/978-3-031-85356-2_9.

[8] P. Mishra, V. Varadharajan, U. Tupakula and E. S. Pilli, "A Detailed Investigation and Analysis of Using Machine Learning Techniques for Intrusion Detection," in IEEE Communications Surveys & Tutorials, 2019 vol. 21, no. 1, pp. 686-728, Firstquarter, doi: 10.1109/COMST.2018.2847722.

[9] Zahid, M., Bharati, T.S. Enhancing cybersecurity in IoT systems: a hybrid deep learning approach for real-time attack detection. Discov Internet Things 2025 5, 73. https://doi.org/10.1007/s43926-025-00156-y

[10] M. Medwed, V. Nikov, J. Renes, T. Schneider and N. Veshchikov, Cyber Resilience for Self-Monitoring IoT Devices, 2021 IEEE International Conference on Cyber Security and Resilience (CSR), Rhodes, Greece, 2021, pp. 160-167, doi: 10.1109/CSR51186.2021.9527995. keywords: {Virtual machine monitors;Computer architecture;NIST;Hardware;Production facilities;Security;Proposals;Cyber resilience;Internet of Things;Device recovery;Trusted services;Self-monitoring;Attack detection},

[11] Shan, L. (IoT) Network intrusion detection system using optimization algorithms. Sci Rep 2025 15, 21706. https://doi.org/10.1038/s41598-025-04638-5

[12] Anjum, M., Dutta, A.K., Elrashidi, A. et al. GraphFedAI framework for DDoS attack detection in IoT systems using federated learning and graph based artificial intelligence. Sci Rep 2025 15, 28050. https://doi.org/10.1038/s41598-025-10826-0

[13] Essop, I.; Ribeiro, J.C.; Papaioannou, M.; Zachos, G.; Mantas, G.; Rodriguez, J. Generating Datasets for Anomaly-Based Intrusion Detection Systems in IoT and Industrial IoT Networks. Sensors 2021, 21, 1528. https://doi.org/10.3390/s21041528

[14] E. Ghiasvand, S. Ray, S. Iqbal, S. Dadkhah, A. Ghorbani. "Resilience Against APTs: A Provenance-Based IIoT Dataset for Cybersecurity Research".

[15] A. Alsaedi, N. Moustafa, Z. Tari, A. Mahmood and A. Anwar, TON_IoT Telemetry Dataset: A New Generation Dataset of IoT and IIoT for Data-Driven Intrusion Detection Systems, in IEEE Access, 2020 vol. 8, pp. 165130-165150, , doi: 10.1109/ACCESS.2020.3022862.

[16] M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan, and R. Jain. WUSTL-IIOT-2021 Dataset for IIoT Cybersecurity Research, Washington University in St. Louis, USA, October 2021

[17] Mohamed Amine Ferrag, Othmane Friha, Djallel Hamouda, Leandros Maglaras, Helge Janicke, January "Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications: Centralized and Federated Learning, IEEE Dataport, 2022 17, doi: https://dx.doi.org/10.21227/mbc1-1h68.

[18] E. Ghiasvand, S. Ray, S. Iqbal, S. Dadkhah, and A. A. Ghorbani. CICAPT-IIOT: A provenance-based APT attack dataset for IIoT environment, preprint, July 2024.

[19] T. Dhaouadi, H. Mrabet, A. Alhomoud and A. Jemai, An Intrusion Detection System Based on HiTar-2024 Dataset Genera-tion from LOG Files for Smart Industrial Internet-of-Things Environment, Computers, Materials & Continua 2025, Vol. 82, Issue 3,

[20] Firouzi, A.; Dadkhah, S.; Maret, S.A.; Ghorbani, A.A. "DataSense: A Real-Time Sensor-Based Benchmark Dataset for Attack Analysis in IIoT with Multi-Objective Feature Selection." Electronics 2025, 14, 4095

[21] S. Tamy, H. Belhadaoui, N. Rabbah, M. Rifi, Ensemble Learning Based Feature Reduction and Selection Methods For Network Intrusion Detection System, Journal of Theoretical and Applied Information Technology 15th July 2021 - Vol. 99. No. 13 – 2021

[22] Nasir Majeed Mir, Sarfraz Khan, Muheet Ahmed Butt and Majid Zaman, An Experimental Evaluation of Bayesian Classifiers Applied to Intrusion Detection, Indian Journal of Science and Technology, 2016, Volume: 9, Issue: 12, Pages: 1-7

[23] Eid, A.M., Soudan, B., Nassif, A.B. et al. Comparative study of ML models for IIoT intrusion detection: impact of data preprocessing and balancing. Neural Comput & Applic 2024 36, 6955–6972. https://doi.org/10.1007/s00521-024-09439-x

[24] T. Dhaouadi, H. Mrabet and A. Jemai, The HiTar-23 Dataset Construction and Validation For Securing Industrial Internet of Things Environment, 2024 IEEE 27th International Symposium on Real-Time Distributed Computing (ISORC), Tunis, Tunisia, 2024, pp. 1-6, doi: 10.1109/ISORC61049.2024.10551372