

# CIC-IIoT-2025 Security Analysis

## Machine Learning for Intrusion Detection in Industrial IoT

Alexis Le Trung, Yahya Ahachim, Rayan Drissi, Aniss Outaleb

ML Security – EPITA SCIA 2026

January 2026

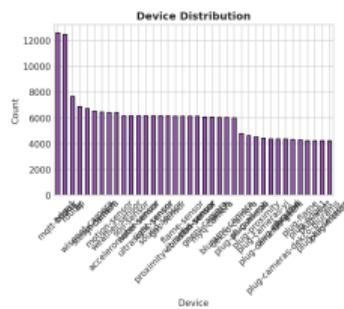
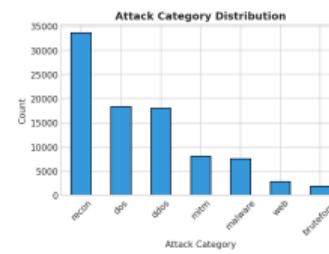
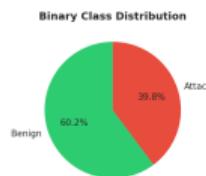
# Agenda

- ① Dataset Overview and Exploration
- ② Anomaly Detection (Unsupervised)
- ③ Classification (Supervised)
- ④ Adversarial Machine Learning
- ⑤ Recommendations

**Objective:** Evaluate ML methods for IIoT intrusion detection and assess adversarial robustness

# CIC-IoT-2025 Dataset

Attribute	Value
Total Samples	227,191
Features	94
Attack Samples	90,391 (39.8%)
Benign Samples	136,800 (60.2%)
Attack Categories	7



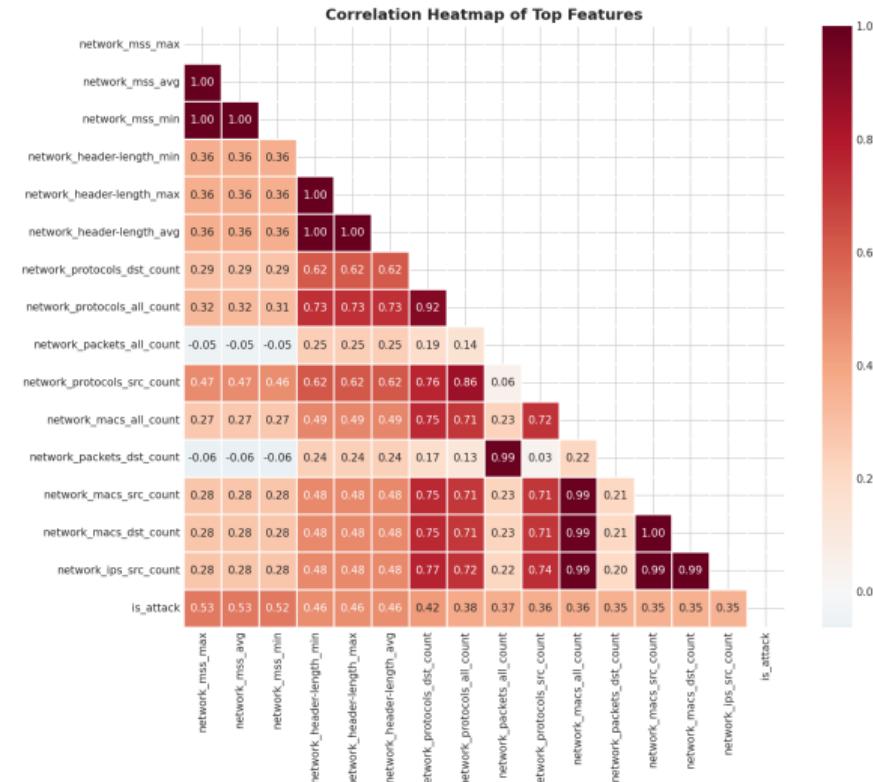
**Categories:** Reconnaissance, DoS, DDoS, MitM, Malware, Web, Brute Force

# Key Discriminative Features

## Top Correlated Features:

Feature	Corr.
network_mss_max	0.526
network_mss_avg	0.525
network_header-length_min	0.464
network_protocols_dst_count	0.423
network_packets_all_count	0.367

TCP MSS and protocol diversity are strong attack indicators



# Anomaly Detection (Unsupervised)

**Trained on benign traffic only** – Detects zero-day attacks

Method	Approach
Isolation Forest	Tree-based isolation via random partitioning
One-Class SVM	Kernel-based boundary in feature space
Local Outlier Factor	Local density deviation detection

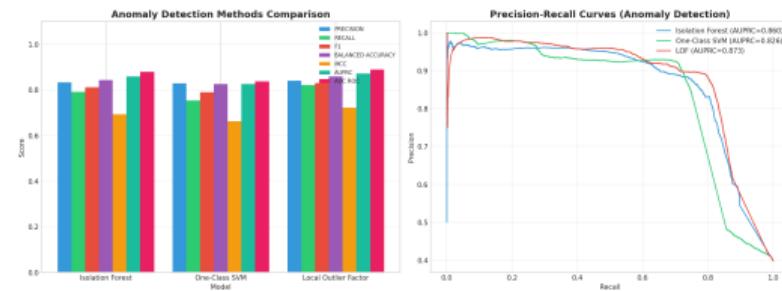
**Evaluation Metric:** AUPRC (Area Under Precision-Recall Curve) – robust for imbalanced detection

# Anomaly Detection Results

Model	F1	AUPRC	MCC
Isolation Forest	0.812	0.860	0.694
One-Class SVM	0.789	0.826	0.663
<b>LOF</b>	<b>0.831</b>	<b>0.873</b>	<b>0.721</b>

**Winner: Local Outlier Factor**

Density-based methods excel on this dataset



# Classification (Supervised)

## Using labeled attack and benign samples

Method	Approach
Random Forest	Ensemble of decision trees with majority voting
Gradient Boosting	Sequential boosting with error correction
SVM (RBF Kernel)	Kernel-based non-linear separation

## Terminology:

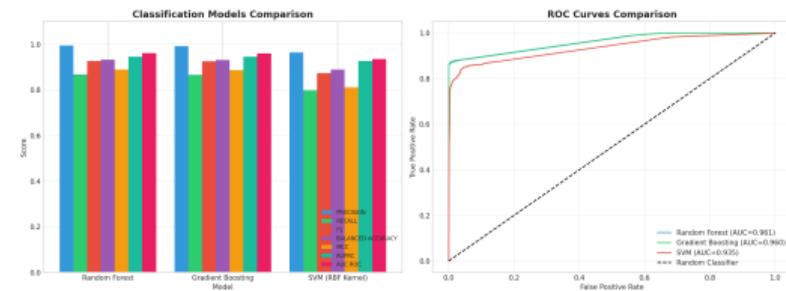
- **Astute Accuracy:** Performance on clean (non-adversarial) data
- **Robust Accuracy:** Performance under adversarial attack

# Classification Results

Model	F1	MCC	AUC
<b>Random Forest</b>	<b>0.927</b>	<b>0.890</b>	<b>0.961</b>
Gradient Boosting	0.925	0.886	0.961
SVM (RBF)	0.874	0.811	0.935

**Winner: Random Forest**

All classifiers achieve >87% F1-score



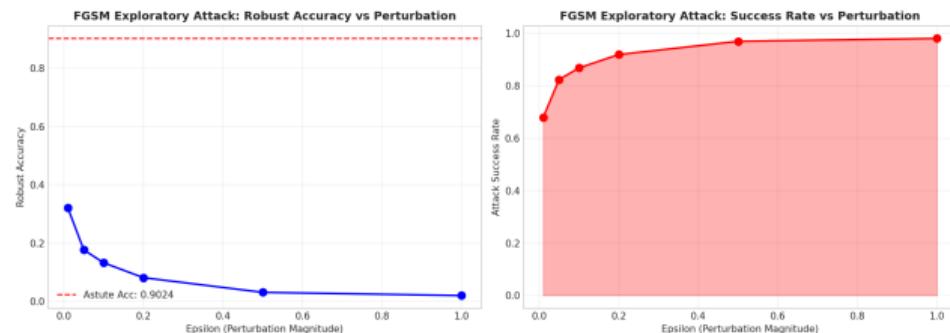
# Adversarial ML: Exploratory Attack (FGSM)

**Fast Gradient Sign Method** – Perturbs test samples

$$x_{adv} = x + \epsilon \cdot \text{sign}(\nabla_x J(\theta, x, y))$$

$\epsilon$	<b>Robust Acc.</b>
0.01	86.7%
0.05	26.0%
0.10	16.0%
0.50	3.5%

Linear SVM (Astute: 89.8%)



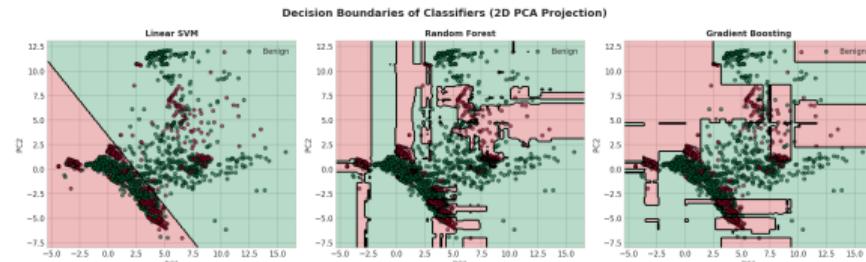
# Adversarial ML: Causative Attack (Data Poisoning)

**Label Flipping** – Poisons training data to shift decision boundary

**Attack Mechanism:**

- Attacker corrupts training labels
- Model learns incorrect boundaries
- Attacks evade detection at inference

Poison Rate	Accuracy
0%	94.3%
10%	85.2%
20%	72.1%



## Model Robustness Comparison ( $\epsilon = 0.5$ )

Model	Astute Acc.	Robust Acc.	Robustness Ratio
Linear SVM	90.2%	3.1%	3.5%
Gradient Boosting	94.4%	34.2%	36.2%
<b>Random Forest</b>	<b>94.6%</b>	<b>41.8%</b>	<b>44.2%</b>

**Finding:** Random Forest retains 44% accuracy under FGSM attack  
Linear models collapse to near-random performance

## Summary: Best Models by Task

Task	Best Model	Key Metric
Zero-day Detection	Local Outlier Factor	$F1 = 0.831$ , $AUPRC = 0.873$
Attack Classification	Random Forest	$F1 = 0.927$ , $AUPRC = 0.946$
Adversarial Robustness	Random Forest	44.2% robust acc.

**Key Insight:** No single model excels at all tasks – defense-in-depth required

# Recommendations

## Multi-Layer Defense Architecture:

- ① **Layer 1 (LOF):** Zero-day attack early warning
- ② **Layer 2 (Random Forest):** Classification with best accuracy and adversarial robustness
- ③ **Layer 3:** Input validation and adversarial training

## Production Hardening:

- Implement adversarial training with augmented samples
- Regular model retraining with new threat intelligence
- Feature monitoring for distribution drift

## Questions?

**CIC-IIoT-2025 Security Analysis**  
Machine Learning for Intrusion Detection

ML Security – EPITA SCIA 2026