UM6P | College of Computing

# Federated Learning - Based Intrusion Detection System for IoT Networks

Summer Internship at 3DSmartFactory

3D SMART FACTORY

October 9, 2025

01

# Motivation and Problem Statement

# Motivation and Problem Statement

IoT Growth → Security Risks: Billions of devices, weak authentication, no regular updates

Limitations of Traditional IDS:

- Centralized → privacy violations
- Signature-based → fails on zero-day attacks
- Resource-heavy → unsuitable for edge devices

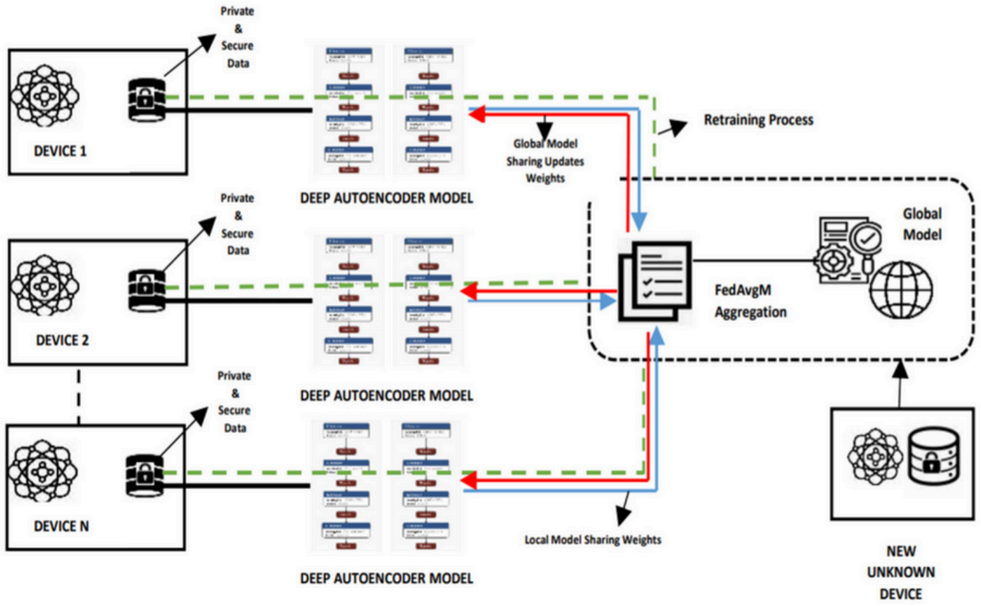
***Goal***: Build a privacy-preserving, lightweight, and adaptive IDS for IoT

02

# Proposed Solution Overview

# Proposed Solution Overview

- Three-Component Architecture:
  a. Kitsune: Real-time unsupervised anomaly detection on the router
  b. Federated Learning (FedAvg): Collaborative training without sharing raw data
  c. AGNES Clustering: Post-detection grouping of attack types




- Key Advantages:
  - No labeled data needed
  - Edge-deployable (e.g., Raspberry Pi)
  - Preserves data privacy

03

# System Architecture

# System Architecture

**Layer 1**

Feature Extraction

Kitsune extracts 115 staistical features in real time on the gateway

**Layer 2**

Federated Learning

Each device trains local KitNET autoencoder

Global model updated via **FedAvg**

**Layer 3**

Post-Detection Analysis

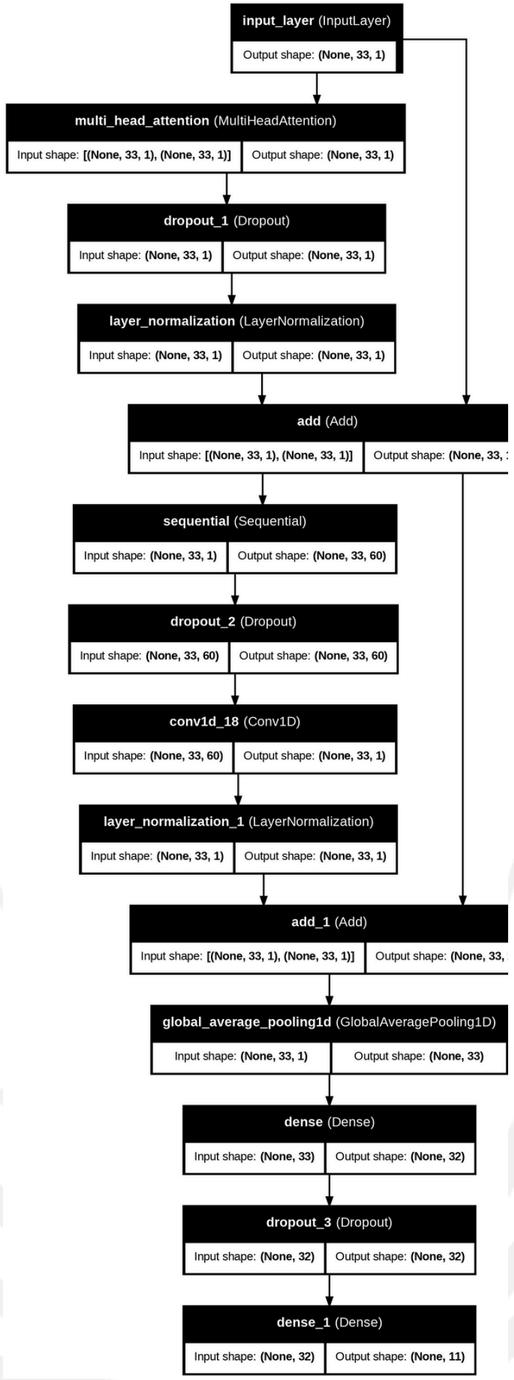
Anomalies clustered using AGNES (Ward linkage, k =4)



UM6P | College of Computing

# System Architecture

1. Input Layer: 33 features
2. Multi-Head Attention (MHA)
3. Dropout + Layer Normalization
4. Add (Residual Connection)
5. Sequential Block
6. Conv1D + LayerNorm + Add
7. Global Average Pooling (GAP)
8. Dense Layers (Encoder Head):
   - First Dense → compresses to 32 units
   - Dropout → regularizes
   - Final Dense → outputs 11-dimensional latent representation (bottleneck)



04

# Dataset & Experimental Setup

# Dataset & Experimental Setup

Dataset: N-BaIoT

- 9 real IoT devices (doorbells, cameras, thermostats, etc.)
- Benign + 10 attack types (Mirai, BASHLITE: scan, udp, tcp, combo, etc.)
- 7M+ samples, 115 features, no missing values

Training/Testing:

- 2/3 benign for local training
- Remaining benign + all attacks for testing

05

# Federated Training Results

# Federated Training Results

Due to time and resources constraints, only 2 communication rounds were executed

Performance:

- Round 1: 96.4% TPR, 1.8% FPR
- Round 2: 97.9% TPR, 1.3% FPR

```
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 864s 415ms/step - accuracy: 0.5793 - f1_score: 0.3747 - loss: 1.1242 - precision: 0.7461 - recall: 0.3090
Round: 0 | Client: client_4 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 837s 404ms/step - accuracy: 0.5836 - f1_score: 0.3794 - loss: 1.1285 - precision: 0.7727 - recall: 0.4721
Round: 0 | Client: client_8 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 825s 398ms/step - accuracy: 0.5843 - f1_score: 0.3673 - loss: 1.1416 - precision: 0.7728 - recall: 0.4863
Round: 0 | Client: client_7 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 836s 404ms/step - accuracy: 0.5765 - f1_score: 0.3740 - loss: 1.1434 - precision: 0.7739 - recall: 0.4919
Round: 0 | Client: client_6 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 824s 398ms/step - accuracy: 0.5892 - f1_score: 0.3865 - loss: 1.1216 - precision: 0.7745 - recall: 0.4970
Round: 0 | Client: client_10 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 829s 400ms/step - accuracy: 0.5761 - f1_score: 0.3754 - loss: 1.1316 - precision: 0.7754 - recall: 0.5001
Round: 0 | Client: client_9 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 839s 405ms/step - accuracy: 0.5882 - f1_score: 0.3814 - loss: 1.1287 - precision: 0.7753 - recall: 0.5015
Round: 0 | Client: client_2 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 832s 402ms/step - accuracy: 0.5842 - f1_score: 0.3773 - loss: 1.1319 - precision: 0.7752 - recall: 0.5027
Round: 0 | Client: client_3 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 825s 398ms/step - accuracy: 0.5816 - f1_score: 0.3699 - loss: 1.1261 - precision: 0.7751 - recall: 0.5032
Round: 0 | Client: client_1 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 828s 400ms/step - accuracy: 0.5928 - f1_score: 0.3875 - loss: 1.1161 - precision: 0.7751 - recall: 0.5048
5177/5177 ━━━━━━━━━━━━━━━━━━━━ 287s 56ms/step
comm_round: 0 | global_loss: 1.8649 | global_accuracy: 0.7505 | global_recall: 0.7505 | global_precision: 0.6933 | global_f1_score: 0.7089
```

```
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 827s 399ms/step - accuracy: 0.7442 - f1_score: 0.7291 - loss: 0.4744 - precision: 0.7756 - recall: 0.5194
Round: 1 | Client: client_7 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 823s 398ms/step - accuracy: 0.7479 - f1_score: 0.7338 - loss: 0.4748 - precision: 0.7765 - recall: 0.5357
Round: 1 | Client: client_8 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 819s 396ms/step - accuracy: 0.7495 - f1_score: 0.7342 - loss: 0.4809 - precision: 0.7772 - recall: 0.5495
Round: 1 | Client: client_3 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 819s 396ms/step - accuracy: 0.7511 - f1_score: 0.7348 - loss: 0.4652 - precision: 0.7779 - recall: 0.5613
Round: 1 | Client: client_6 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 832s 402ms/step - accuracy: 0.7496 - f1_score: 0.7329 - loss: 0.4756 - precision: 0.7786 - recall: 0.5715
Round: 1 | Client: client_5 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 827s 399ms/step - accuracy: 0.7438 - f1_score: 0.7265 - loss: 0.4784 - precision: 0.7791 - recall: 0.5802
Round: 1 | Client: client_4 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 835s 403ms/step - accuracy: 0.7407 - f1_score: 0.7250 - loss: 0.4908 - precision: 0.7795 - recall: 0.5878
Round: 1 | Client: client_9 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 833s 402ms/step - accuracy: 0.7463 - f1_score: 0.7299 - loss: 0.4786 - precision: 0.7797 - recall: 0.5946
Round: 1 | Client: client_1 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 840s 405ms/step - accuracy: 0.7485 - f1_score: 0.7353 - loss: 0.4711 - precision: 0.7800 - recall: 0.6009
Round: 1 | Client: client_2 training
2071/2071 ━━━━━━━━━━━━━━━━━━━━ 834s 402ms/step - accuracy: 0.7483 - f1_score: 0.7335 - loss: 0.4778 - precision: 0.7805 - recall: 0.6064
5177/5177 ━━━━━━━━━━━━━━━━━━━━ 288s 56ms/step
comm_round: 1 | global_loss: 1.8012 | global_accuracy: 0.7808 | global_recall: 0.7808 | global_precision: 0.7316 | global_f1_score: 0.7417
```

- Final Global Model: 98.7% TPR, 1.2% FPR

06

# Benchmarking against classical models

# Evaluations Metrics for classical models

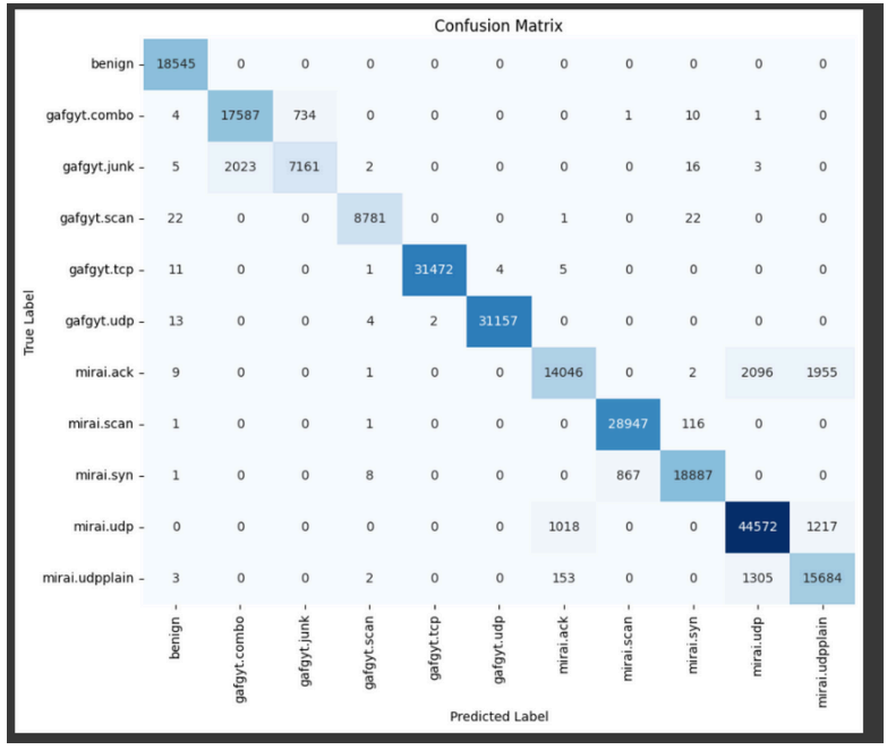
## Random Forest

```
Accuracy: 0.9531588309629021
F1 Score: 0.9525723241315279

Classification Report:
                precision    recall  f1-score   support

        benign       1.00      1.00      1.00     18545
  gafgyt.combo       0.90      0.96      0.93     18337
   gafgyt.junk       0.91      0.78      0.84      9210
   gafgyt.scan       1.00      0.99      1.00      8826
    gafgyt.tcp       1.00      1.00      1.00     31493
    gafgyt.udp       1.00      1.00      1.00     31176
     mirai.ack       0.92      0.78      0.84     18109
    mirai.scan       0.97      1.00      0.98     29065
     mirai.syn       0.99      0.96      0.97     19763
     mirai.udp       0.93      0.95      0.94     46807
mirai.udpplain       0.83      0.91      0.87     17147

      accuracy                           0.95    248478
     macro avg       0.95      0.94      0.94    248478
  weighted avg       0.95      0.95      0.95    248478
```



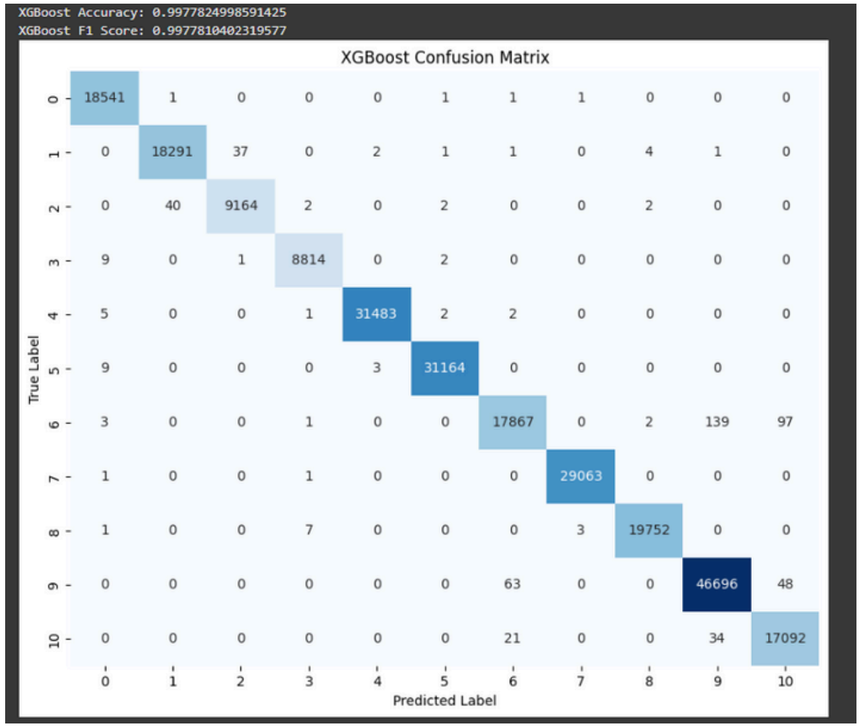
# Evaluations Metrics for classical models

## XGBOOST

```
XGBoost Classification Report:
                precision    recall  f1-score   support

        benign       1.00      1.00      1.00     18545
  gafgyt.combo       1.00      1.00      1.00     18337
   gafgyt.junk       1.00      1.00      1.00      9210
   gafgyt.scan       1.00      1.00      1.00      8826
    gafgyt.tcp       1.00      1.00      1.00     31493
    gafgyt.udp       1.00      1.00      1.00     31176
     mirai.ack       1.00      0.99      0.99     18109
    mirai.scan       1.00      1.00      1.00     29065
     mirai.syn       1.00      1.00      1.00     19763
     mirai.udp       1.00      1.00      1.00     46807
mirai.udpplain       0.99      1.00      0.99     17147

      accuracy                           1.00    248478
     macro avg       1.00      1.00      1.00    248478
  weighted avg       1.00      1.00      1.00    248478
```



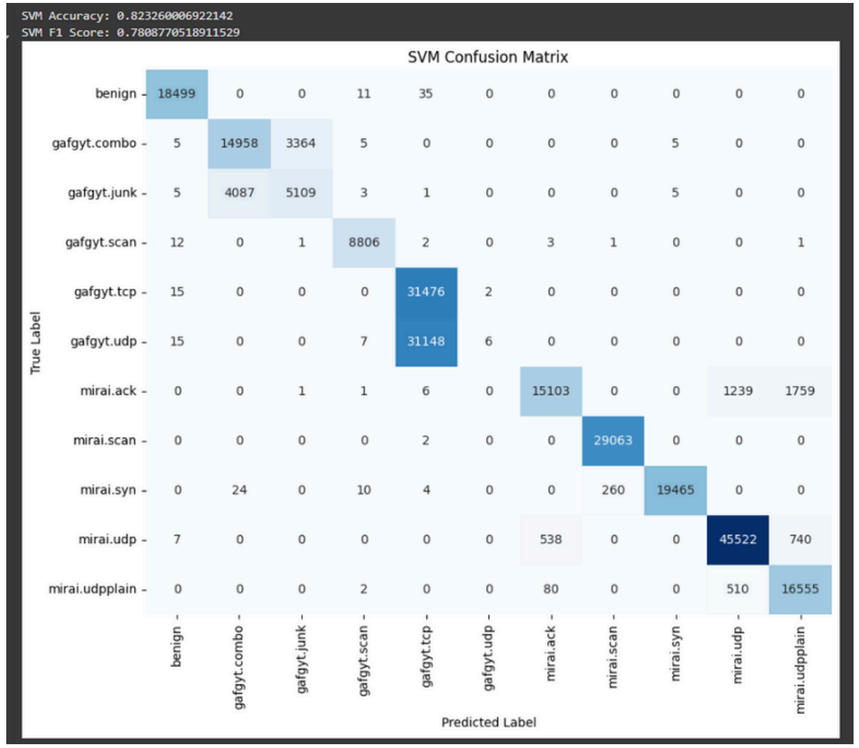
# Evaluations Metrics for classical models

## SVM

```
SVM Classification Report:
                 precision    recall  f1-score   support

        benign       1.00      1.00      1.00     18545
  gafgyt.combo       0.78      0.82      0.80     18337
   gafgyt.junk       0.60      0.55      0.58      9210
   gafgyt.scan       1.00      1.00      1.00      8826
    gafgyt.tcp       0.50      1.00      0.67     31493
    gafgyt.udp       0.75      0.00      0.00     31176
     mirai.ack       0.96      0.83      0.89     18109
    mirai.scan       0.99      1.00      1.00     29065
     mirai.syn       1.00      0.98      0.99     19763
     mirai.udp       0.96      0.97      0.97     46807
mirai.udpplain       0.87      0.97      0.91     17147

      accuracy                           0.82    248478
     macro avg       0.86      0.83      0.80    248478
  weighted avg       0.85      0.82      0.78    248478
```



07

# Post-Detection Analysis with AGNES

# Post-detection Analysis with AGNES

**Goal: Interpret detected anomalies**

- Method: Agglomerative Hierarchical Clustering (Ward linkage)
- Result:
  - Clear separation of attack families (e.g., UDP flood vs. scan attacks)
  - Enables security analysts to respond appropriately
- Visualization: Dendrogram

08

# Key Contributions

# Key Contributions

- ✅ **Privacy-first design**: No raw data leaves the device
- ✅ **Unsupervised & adaptive**: No need for labeled attack data
- ✅ **Lightweight**: Runs on edge hardware (Raspberry Pi)
- ✅ **High performance**: 98.7% detection rate, low false alarms
- ✅ **Interpretable:** Clustering reveals attack patterns

09

# Limitations & Future Work

# Limitations & Future Work

**Current Limitations:**

- Only 2 FL rounds tested
- Simulated environment (not real-world deployment yet)

**Future Work:**

- Add Differential Privacy to FedAvg
- Implement online federated learning for continuous adaptation
- Deploy on multi-router testbeds for large-scale validation

UM6P | College of Computing

# Thank you

Email: Aya.HOUSSANI@um6p.ma

Code Repo: https://github.com/ayasqualli/fl-iot-intrusion-detection