

# Non-IID Data Distribution Impacts on Federated Intrusion Detection Model Convergence and Accuracy

Assignee Research

May 31, 2026

## Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: What is the effect of non-IID data distribution across federated clients on the convergence rate and final accuracy of intrusion detection models trained on the Edge-IIoTset dataset. This paper presents a large-scale empirical study aimed at identifying the optimal local deep learning model and data volume for deploying intrusion detection systems (IDS) on resource-constrained IoT devices using federated learning (FL). While previous studies on FL-based IDS, 4 claims were extracted from source literature; 4 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 8.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: Federated Learning-Based Intrusion Detection in IoT Networks: Performance Evaluation and Data Scaling Study. Research question: What is the effect of non-IID data distribution across federated clients on the convergence rate and final accuracy of intrusion detection models trained on the Edge-IIoTset dataset?.

## 2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 8.3/10.

### 3 Results

11 papers retrieved. 4 claims extracted; 4 independently verified. Quality review score: 8.3/10.

### 4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

### 5 Extracted Claims

Claim	Verified	Confidence
The study evaluates how detection accuracy, convergence speed, and inference costs (latency and model size) vary across	✓	0.33
CNN achieves the best trade-off between detection performance and computational efficiency, reaching ~98% accuracy with	✓	0.32
The more complex CNN+BiLSTM architecture yields slightly higher accuracy (~99%) at a significantly greater computational	✓	0.30
Deployment tests on Raspberry Pi 5 devices confirm that all three models can be effectively implemented on real-world Io	✓	0.31

### References

- <https://doi.org/10.3390/network3010008>
- <https://doi.org/10.3390/jsan14040078>
- <https://doi.org/10.1109/access.2022.3165809>