# Faulty State Classification in Cellular Networks under Label Scarcity: Classical vs. Few-Shot Learning Using KPI Data

**Riset Unggulan ITB**

Hajiar Yuliana, Iskandar, Hendrawan

STEI ITB

## Abstract

Reliable classification of faulty operational states in cellular networks is challenging due to limited labeled anomaly data. This study evaluates Prototypical Networks for KPI-based fault classification using a real-world 4G LTE dataset from Cianjur, Indonesia, covering four anomaly types: cell outage, sleep cell, degraded performance, and handover failure. Two embedding architectures, dual-layer Long Short-Term Memory (LSTM) and hybrid Autoencoder-LSTM (AE+LSTM), are compared across 1-shot to 10-shot configurations over 20 runs. The LSTM-based model achieves F1-scores of 0.9268 at 1-shot and 0.9808 at 10-shot, with ROC-AUC of 0.9738. In contrast, AE+LSTM performs considerably worse, indicating that autoencoder compression may discard temporal features essential for anomaly discrimination. Results demonstrate the effectiveness of LSTM-based temporal embeddings for few-shot fault classification, requiring only ten labeled samples per class for reliable classification of sufficiently represented anomalies. Degraded performance remains a boundary case due to only six instances from a single site.

Keyword: few-shot learning, prototypical networks, KPI-based operational state classification, cellular networks, LSTM embedding, label efficiency, shot sensitivity, 4G LTE.

## Introduction

Cellular network fault classification using Key Performance Indicators (KPIs) is challenging due to the scarcity of labeled anomalies and their complex temporal characteristics. Conventional supervised learning methods often require extensive labeled datasets, making them less suitable for operational environments with limited fault records. Few-shot learning (FSL), particularly Prototypical Networks, offers an alternative by enabling classification from limited labeled examples. However, the effects of embedding architecture and the minimum number of labeled samples required for reliable classification remain insufficiently explored. This study addresses these gaps by comparing LSTM and hybrid Autoencoder-LSTM embeddings across different shot configurations using real-world 4G LTE KPI data from Cianjur, Indonesia.

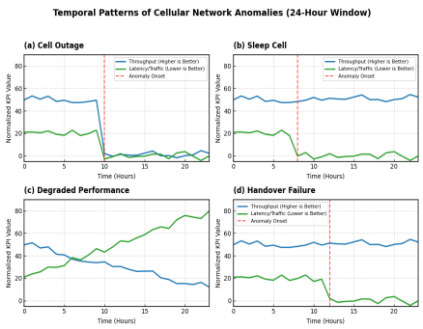



*Figure 1. Temporal Patterns of Cellular Network Anomalies Over 24-Hour Observation Window*

## Research Methodology

This study employs Prototypical Networks to investigate embedding architecture effects and minimum labeled sample requirements for cellular network fault classification. The dataset comprises 70,779 non-overlapping 24-hour KPI windows, each containing 20 features collected from operational 4G LTE sites in Cianjur, Indonesia, between December 2024 and February 2025. Four anomaly types are identified: cell outage, sleep cell, degraded performance, and handover failure. Data preprocessing includes cleaning, mean imputation, outlier inspection, correlation analysis, and StandardScaler normalization. No synthetic augmentation is applied to few-shot experiments. Chronological separation prevents temporal data leakage. Prototypical Networks are evaluated using 2-way 1-, 3-, 5-, 7-, and 10-shot configurations, comparing dual-layer LSTM and hybrid AE+LSTM embeddings. Each model undergoes 1,000 training episodes and 20 independent experimental runs. Degraded performance is excluded from K-shot evaluation due to insufficient samples. Performance is assessed using accuracy, precision, recall, F1-score, ROC-AUC, confusion matrices, and false alarm rate (FAR).

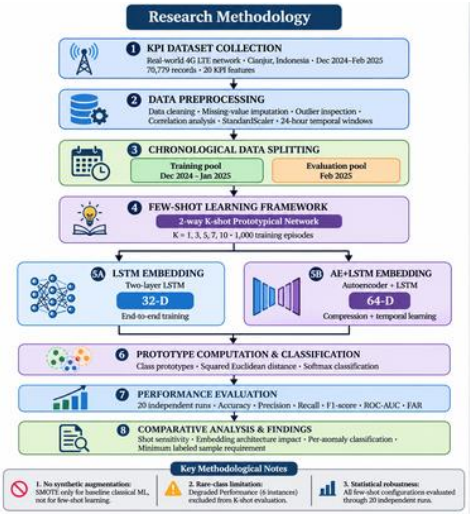



*Figure 2. Research methodology*

## Simulation Result

This simulation result discuses four classical supervised models—Random Forest (RF), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and XGBoost—using the SMOTE-balanced dataset. Their results establish a reference performance level under synthetic data augmentation and illustrate classical models' dependence on labeled data, motivating few-shot learning approaches examined in this results.

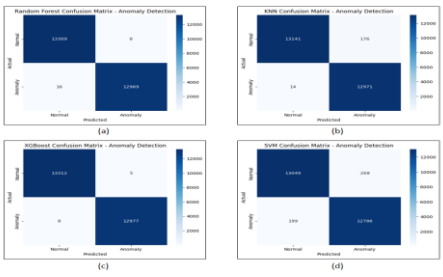



*Figure 3. Confusion matrices of classical ML models trained on the full SMOTE-balanced KPI dataset*

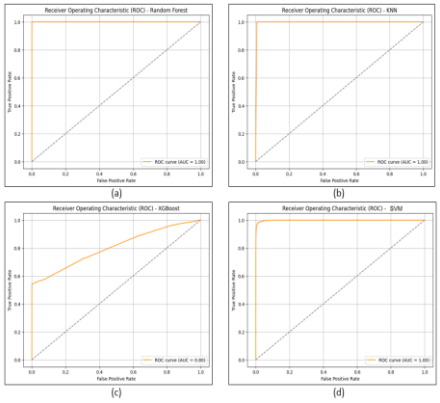



*Figure 4. ROC curves of classical ML models trained on the full SMOTE-balanced KPI dataset*

## Conclusion and Discussion

1. **Attention-Weighted Prototypes**
   Improve prototype quality and robustness to noisy labels by assigning adaptive weights to representative support samples.
2. **Online Prototype Adaptation**
   Develop incremental prototype updates using newly verified anomaly data from NOC operations without requiring full model retraining.
3. **Transformer-Based Temporal Embedding**
   Investigate Transformer architectures with multi-head self-attention to capture long-range temporal dependencies in KPI sequences and potentially improve classification performance.
4. **Cross-Site and Multi-Operator Validation**
   Evaluate model generalization across different sites, operators, traffic patterns, and network configurations, including the robustness of the K = 10 sample requirement.
5. **Cross-Site Data Pooling for Rare Anomalies**
   Aggregate verified anomaly samples across multiple sites or operators to enable reliable K-shot evaluation of extremely rare cases, particularly Degraded Performance.