High-Level Design Document

Project: AI/GenAI-based Telecom Log Analysis
Version: Baseline v1

---

1. Introduction

Telecom networks generate large volumes of operational and error logs. Traditional log analysis is rule-based and requires domain experts. This project introduces an AI-driven solution that leverages **pretrained Transformer models** to classify and summarize logs, enabling faster issue triaging and improved observability.

---

2. Objectives

- Automate classification of log messages into categories (e.g., error, warning, config issue).
- Generate **summaries of raw log data** using Generative AI.
- Establish a baseline architecture that can later be extended for explanations, anomaly detection, and remediation suggestions.

---

3. System Overview

Components:

1. **Log Input Module**
    - Source: Raw 5G/telecom logs (text-based).
    - Current implementation: `log_generate_parse.get_raw_logs()`.

2. **AI/ML Layer**
    - **Zero-Shot Classification**:
        - Model: `facebook/bart-large-mnli`.
        - Function: Categorizes log entries without supervised training.
    - **Summarization (GenAI)**:
        - Model: `facebook/bart-large-cnn`.
        - Function: Generates human-readable summaries of multiple log entries.
3. **Output Module**
    - Console-based results.
    - Displays classification categories and generated summaries.

---

4. Key Technologies

- **Python 3.11** (baseline tested)
- **Hugging Face Transformers** (`pipeline`)
- **PyTorch** backend
- **Models**:
    - `facebook/bart-large-mnli` → classification
    - `facebook/bart-large-cnn` → summarization

5. Next Steps (Future Extensions)

- Add **explanation generation** (why the log occurred, plain-English causes).

- Integrate **anomaly detection** with statistical/ML methods.
- Support **multi-vendor log formats**.
- Wrap in **FastAPI/Streamlit** for interactive use.
- Extend to **deployment in enterprise observability systems**.