Short Report: Telecom KPI Anomaly Detection Agent

Project Overview:

This project integrates an unsupervised anomaly detection pipeline with a conversational AI agent to monitor and analyze telecom KPI trends across multiple network sites and sectors. The system allows users to query anomalies interactively and receive insightful, data-backed responses.

Approach & Architecture:

Data: Daily KPI data collected from multiple sectors/sites across a telecom network, including metrics like RSRP, SINR, Throughput, Call Drop Rate, RTT, and more.

Preprocessing:

- Normalized dates using pandas.to_datetime.
- Enforced domain-informed KPI bounds to eliminate implausible values and reduce noise.
- Filtered out sectors with fewer than 10 valid records for a KPI.

Models Used:

DWT-MLEAD: Detects frequency-based anomalies such as seasonal shifts and shape distortions in time series.

Isolation Forest: Tree-based detector for outliers in multivariate high-dimensional data.

Ensemble Voting: Combines both methods to reduce false positives and improve robustness.

Agent Architecture:

- Built using **LangChain**, **LangGraph**, and **NVIDIA NeMo LLMs**.
- Integrates multiple tools for anomaly insights, extreme value detection, KPI comparison, and Granger causality.
- Uses **FastAPI** (MCP_server.py) as the backend, and **Gradio UI** (app.py) for chat interface.

Design Decisions:

- **Model Choice**: Selected complementary models (DWT-MLEAD and Isolation Forest) based on robustness to time-series patterns and outliers. Ensemble voting improves confidence in detected anomalies.
- **Free-tier API Usage**: Used NVIDIA NIM and Tavily Search APIs due to cost efficiency while maintaining functionality during prototyping.
- **Streaming Agent Workflow**: Deployed with streaming support for real-time responses and continuous tool invocation using LangGraph conditional loops.

Key Challenges & Solutions:

Challenge	Solution
Lack of labeled anomalies	Employed visual verification and anomaly rate heuristics (0.5%–5%)
Noise and extreme outliers	Applied domain-informed bounds (e.g., RSRP ∈ [-120, -60 dBm])
Temporal inconsistency	Grouped and sorted time-series by Sector_ID before modeling
Tool integration with agent	Used LangChain tools and conditional flow to trigger tool usage automatically
Multi-KPI anomaly reasoning	Designed tools to compare related KPIs and co-occurrence patterns

Example Interactions:

User Query	Assistant Response Summary
"Were there anomalies in DL_Throughput?"	Yes, 12 anomalies detected. Avg anomaly above baseline = 35.2 Mbps. Most occurred between June 15–21. Related SINR anomalies suggest signal degradation.

User Query	Assistant Response Summary
“What’s the worst day for Call Drop Rate?”	July 2nd had the peak with 14 anomalies. Average Call Drop Rate = 3.2%. Possibly due to poor handovers.
“Which site had the lowest SINR last week?”	Site_047 showed lowest SINR average of -17.6 dB. Likely in poor coverage area.

References:

- **Schmidl, S., Wenig, P., & Papenbrock, T. (2022).** *Anomaly Detection in Time Series: A Comprehensive Evaluation*
  Hasso Plattner Institute, University of Potsdam; Philipps University of Marburg
- LangChain, NVIDIA NeMo Instruct Models, Tavily Search API, FastAPI, Gradio