Name: Rajneesh Tiwari

Email: Rajneesh.vish1@gmail.com

Affiliation: Optum AI Pvt Ltd

1. Approach Overview

Our solution uses a two-stage fine-tuning pipeline: reasoning trace generation followed by Supervised Fine-Tuning (SFT) and Group Relative Policy Optimization (GRPO) on Qwen3-32B. The core insight is that 5G troubleshooting follows well-defined engineering rules - rather than relying on a general-purpose LLM to reason from raw data, we encode expert domain knowledge into structured reasoning traces, then train the LLM to reproduce and generalize that reasoning.

2. Reasoning Trace Generation

Question Taxonomy & Reasoning Trace generation

The system segments questions into three distinct categories: Type A (cell parameter analysis), Type B (drive test root cause analysis), and Generic queries.

Analytical logic involves a 5-tier hierarchical classification system for Type A metrics and an 8-rule cascade for Type B that integrates multiple data tables and PHY-layer health checks.

Training relies on data analysis driven structured Chain-of-Thought traces generated with 98.3% rule accuracy using training data, supplementing the few failures with "expert correction" branches to teach edge-case reasoning.

To ensure data ingestion stability, the architecture uses a header-based table parser that remains robust to column reordering across variable question formats.

3. SFT & GRPO Training

We fine-tuned Qwen3-32B (4-bit quantized via bitsandbytes NF4) using Unsloth's QLoRA implementation: -

- Training data: 2,400 Type A reasoning traces generated from the labeled training set using validated mathematical rules.

The SFT model learns the structured reasoning patterns from the traces. The GRPO model builds on top of trained SFT model for 100 steps and finetunes the model even more for format, accuracy, and reasoning based rewards.

4. Data Privacy and Compliance

All training data comes exclusively from the competition dataset provided by Zindi. No external personal data, user information, or proprietary telecom subscriber data was used. The 5G network metrics in the dataset are anonymized cell-level measurements (RSRP, SINR, throughput, handover counts) that contain no personally identifiable information. Our model processes only aggregated network performance metrics and cannot identify individual users or devices.

5. Model Security Risks

The model is specialized for 5G root cause classification - a narrow, well defined task. It outputs only structured labels (C1-C8, A-I, or numeric answers) wrapped in a fixed format. The system prompt constrains outputs to the rule framework, limiting the attack surface for prompt injection. However, as with any LLM, adversarial inputs could potentially cause misclassification. In a production deployment, both inputs and outputs should be validated against physically plausible ranges before acting on them.

6. Data and Model Access Control

The SFT and GRPO models are hosted on Hugging Face Hub (Phaedrus33/SFT_final_submission, Phaedrus33/GRPO_final_submission). The training code and classification pipeline are version-controlled on GitHub. No API keys, tokens, or credentials are stored in the codebase - authentication uses environment variables. The base model (Qwen3-32B) is open-source under Apache 2.0.

7. Edge Computing Considerations

The solution uses a 32B parameter model, which requires substantial GPU resources (1x A100/H200 80-140GB) for inference.

8. Data Governance

The training pipeline is fully reproducible: run_final.sh --all regenerates traces from the competition training CSV, trains SFT, trains GRPO, and produces the submission. Traces are generated exclusively from train.csv using generate_traces_final.py. No test set data is used in trace generation or model training. No data augmentation from external sources