MEDIGUIDE

Problem Statement :

Healthcare providers increasingly seek AI-assisted tools to deliver preliminary guidance and triage for common medical inquiries. A conversational “e-doctor” can offer patients immediate, text-based responses to symptom-related questions, while clarifying that it does not replace professional diagnosis

Transform an open-source, decoder-only Transformer model into a medically informed chatbot capable of responding to patient questions with accurate, contextually appropriate, and professionally worded advice. Compare three distinct fine-tuning strategies in terms of performance, resource utilization, and deployment feasibility.

Goals:

- Ingest a user’s free-text medical question and generate an answer that :
  - Adheres to recognized clinical guidelines.
  - Uses formal, professional language.
  - Includes disclaimers regarding non-substitutive informational use.
- Supports fine-tuning modalities on identical training data like:
  - Prompt Tuning
  - Parameter-Efficient Fine-Tuning using the PEFT library (supports Adapter Training, LoRA, QLoRA and 8-bit quantization)
- Performance Report (PDF) including:
  - Dataset description and splits
  - Comparative results table summarizing ROUGE, PPL, latency, and model size.
  - Summary of trade-offs and recommended deployment strategy
- Taking steps to ensure no patient-identifiable information is present; abide by HIPAA-equivalent anonymization standards.

TechStack/Frameworks :

- Machine Learning : Python, with PyTorch, HugginFace Accelerate
- Transformers & Fine-Tuning : HugginFace Transformers, PEFT, BitsAndBytes

Mentors

Shirshendu
9674397828

Sagnik
9874693499