Post-Hackathon Finetuning Challenge

Deadline: Tuesday (20 May 2025)
Teams: Groups of 2 (same as before)

Goal

Now that you've experimented with building GenAI applications using APIs like Gemini, it's time to go deeper: **learn when and how to finetune models**. You'll be working hands-on with both encoder and decoder models.

This exercise will give you practical insight into:

- When to use **encoder models** (e.g., for classification, retrieval, sentence embeddings)
- When to use **decoder models** (e.g., for generation, summarization, dialogue)
- How to choose the **right datasets** for your task
- How to **finetune and evaluate** models for your needs, not just rely on LLM APIs

Your Task

- **Pick an application idea** — you can loosely base this on what you worked on during the hackathon.
- Break problem into two parts:
  - **Encoder model finetuning** (e.g., BERT or DistilBERT for classification or embeddings).
  - **Decoder model finetuning** (e.g., Qwen, LLaMA for generation tasks).
- Showcase how both approaches could serve a broader application (e.g., encoder for intent detection, decoder for response generation).
- If your idea doesn't map well to both model types, **pivot** or **adapt**.
- **Finding a suitable dataset for finetuning is part of the challenge.**

You do **not** need to build an app. The focus is:

- Successfully finetuning the model
- Showing evaluation results (e.g., classification accuracy, sample generations)
- Explaining the *why* of your choices

Resources

- **Decoder model finetuning code** has been shared with you already.
- **Encoder model finetuning code** was part of your induction challenge — use it as reference.

- You may use Hugging Face datasets or prepare your own synthetic datasets using GPT/Gemini (best of luck generating enough samples 😉)

Deliverables (due Monday)

- Just present your work and findings.

Why This Matters

Large Language Models are not always the right solution — sometimes, a small finetuned model does the job better and cheaper. This challenge is about learning how to **make that decision intelligently** and **execute with confidence**.