Group 4:
Chen Zhiyang
Lakshmi Aishwarya Vathada

PROJECT TITLE: Evaluating the Performance of Distilled Open-Source Models Using DeepSeek R1: A Comparative Study of Pre- and Post-Distillation Performance

Workflow:

Steps:

1. Research on Distillation and Reasoning Models. *(Both)*
2. Application of Distillation using Deepseek R1 on Various Open Source models:
   a. IBM Granite *(Chen)*
   b. LLaMA 3.1 *(Chen)*
   c. Mistral *(Chen)*
   d. Qwen 2.5 *(Lakshmi)*
   e. DBRX *(Lakshmi)*
   f. Microsoft Phi *(Lakshmi)*
3. Evaluate model before and after distillation using *(Both)*:
   a. English
      i. MMLU
      ii. AGIEval
   b. Math
      i. MATH
      ii. MGSM
   c. Code
      i. HumanEval
      ii. APPS
   d. Mandarin
      i. C-Eval
      ii. CLUEWSC
4. Prepare project summary and presentation which includes *(Both)*:
   a. Theory of Distillation and Explanation of "Reasoning model"(Deepseek R1 and OpenAI o1)
   b. Understanding and explaining the different models briefly
   c. Understanding and explaining the evaluation metrics we chose and why
   d. LLMs evaluation before distillation and after distillation
   e. Discussion on results, key insights, conclusions, future research.