Debate-Based Reasoning for Math Questions using Groq

Motivation

Why a Debate-Based Approach?

- Large Language Models (LLMs) sometimes make reasoning errors
- Traditional prompting gives **one-shot** solutions.
- **Debate strategy** uses *critique and revision* to refine answers.

What is Groq?

- Paid Inference API for Fast LLMs
- Provides access to models like:
- `llama-3.1-8b-instant`
- `gemma2-9b-it`
- Known for **ultra-low latency** inference (as low as 1ms/token).
- Useful for **real-time, multi-turn prompting** like debate.

GSM8K Dataset

- **GSM8K = Grade School Math Benchmark**

- Benchmark for math word problems.

- Each question includes:
    - A multi-step math problem
    - A free-text solution ending with a numeric answer

How to Access Groq

- **Steps to Use Groq API**

- Get API Key from https://console.groq.com

- Install the SDK:
  `pip install groq pandas datasets`

- Save your key:
  `groq_api.txt`

- Use `Groq(api_key=...)` in Python