AMD AI SPRINT Track 1

CSA-CDS

Shiva Chandhra
Vishal

Data Generation

- Used ChatGPT with controlled prompts to synthesize Q&A examples

- Prompt template includes: topic tag, multi-speaker statements, 4 MCQ options, step-by-step explanation, correct label

- Manual spot review: sample 10% of generated data for clarity and correctness

- Total Examples: 1500

Training

- Library: trl (Hugging Face's RL & SFT toolkit)

- Base checkpoint: Qwen-3-4B

- Technique: Supervised Fine Tuning

- Hyper Parameters:
  - Learning Rate (lr): 1e-5
  - Num epochs (num_epochs): 25
  - Batch size (batch_size): 16
  - Lora Rank (lora_rank): 32
  - Lora alpha (lora_alpha): 64

Github and Hugging face links

- Github link: https://github.com/shivaiisc/AAIPL_CDS_CSA

- Hugging face link: https://huggingface.co/pvishal/qwen3_4B_aaipl_csa_cds_team

Thank You