

AMD AI Premier League

Reinforcement Learning Hackathon – IIT Delhi

A competitive AI challenge requiring head-to-head agent deployment on AMD MI300 GPUs with constrained local models. Participants develop Q-Agents to generate challenging MCQs and A-Agents to accurately answer opponent questions under strict validation rules.

TEAM NAME - THE DECIDERS



Problem Statement & Competition Objective

Q-Agent Challenge

Generate valid, challenging multiple-choice questions that maximize opponent error rates while maintaining $\geq 50\%$ format correctness under strict JSON schema validation.

A-Agent Task

Maximize correct answer rate on opponent-generated questions through optimized inference strategies, low-temperature sampling, and deterministic response generation.

Constraints

English-only responses, no external APIs or RAG systems allowed, local HuggingFace models only, competitive evaluation with automatic disqualification for format violations.

System Architecture & Training Pipeline

01

Base Model Selection

Qwen-based and LLaMA-based architectures chosen for superior reasoning capabilities on MCQ generation and comprehension tasks

02

Supervised Fine-Tuning

Curated MCQ datasets applied for instruction tuning on question generation and answer recognition patterns

03

Parameter-Efficient Training

LoRA adapters implemented to reduce GPU memory footprint while maintaining model capacity

04

Preference Optimization

Optional DPO (Direct Preference Optimization) applied for reward modeling and response quality improvement

05

AMD MI300 Optimization

bfloat16 precision and gradient checkpointing enabled for maximum throughput on target hardware platform

Evaluation & Scoring Methodology

100%

Correct Answer Rate

Percentage of opponent questions answered correctly by A-Agent

100%

Invalid Question Rate

Percentage of questions causing opponent errors by Q-Agent

50%

Minimum Validity

Strict lower bound for format-correct questions

Validation Requirements

- Exact JSON schema compliance enforced
- MCQ format must include A/B/C/D options
- Automatic disqualification triggered if format-correct questions drop below 50%
- String matching applied for answer verification

Strategic Considerations

- Q-Agent balances difficulty against validation reliability
- A-Agent uses low-temperature sampling for consistent predictions
- Prompt engineering critical for format adherence

Deliverables & Development Artifacts

Required Submission Components

- Trained Q-Agent model weights (adapter-only or full checkpoint)
- Trained A-Agent model weights with inference configuration
- Training scripts including LoRA/DPO implementation details
- Evaluation harness validating format correctness
- Sample question/answer outputs demonstrating system behavior
- This technical documentation for judging panel review

Competition Compliance Summary

Rule Adherence

System fully implements competition requirements including local HuggingFace model restriction, no external API calls, English-only I/O, strict JSON validation, and dual-agent architecture.

Hardware Optimization

Training and inference pipelines optimized for AMD MI300 GPU architecture using bfloat16 precision and memory-efficient LoRA parameter tuning.

Evaluation Ready

Validation methodology matches judging criteria with JSON schema enforcement, MCQ format checking, and automatic disqualification logic for format violations.

THE END.