Exam
Due Mar 18, 10 pm

1. Compare the advantages and disadvantages of fine-tuning a pretrained LLM versus continuous pretraining. Provide examples of how each approach is more beneficial.

2. Explain LoRA (Low-Rank Adaptation) and QLoRA. How do these techniques enable efficient fine-tuning of large models on consumer-grade GPUs? What advantage does QLoRA have over LoRA?

3. Explain the differences between Generalized Reinforcement Policy Optimization (GRPO) and Direct Preference Optimization (DPO).

4. What types of tasks would GRPO be preferable over DPO and vice versa?

5. What is a catastrophic policy shift in DPO? How does KL divergence regularization help prevent it?

6. Define catastrophic forgetting in the context of LLMs. Propose and evaluate strategies to mitigate this issue when fine-tuning an LLM on a domain-specific dataset.

7. Use bitsandbytes to quantize the model you used in assignment two. Compare the memory (CPU and GPU) and run time required by a FP16 quantized model, an 8-bit quantized model, and a 4-bit model. How does the 8-bit model output compare with the full model?

8. Train GPT-2 model for code generation using DPO and the MBPP (Mostly Basic Python Problems) data. How well does the model perform in code generation?

9. Repeat #8 but use unsloth and vllm in the model's training. Compare the memory (CPU and GPU) and run time required without using unsloth and vllm (#8) with using them (this question).

10. (Extra Credit) Train GPT-2 model for code generation using GRPO with unsloth and VLLM. How does the model's performance compare with the model from #8?

Larger datasets can be used to train models in 8 through 10, which will produce better results but require longer training times.

What to turn in.

Answers to questions 1 through 6 can be turned in as a PDF or a Word Document. Answers to questions 7 through 10 are to be done in a Jupyter notebook in the style of assignment 2v2. Combine all documents in a zip file to turn in the exam on Canvas.

You will turn in the assignment in Canvas.

Grading

	Percent of Grade
Each Problem	10 points

The exam is worth 90 points. If you answer the extra-credit question, you can earn 100 points out of 90.

Late Policy

An assignment turned in 1-7 days late, will lose 10% of the total value of the assignment per day late. Once a solution to an assignment has been posted or discussed in class, the assignment will no longer be accepted. Late penalties are always rounded up to the next integer value.