EEP 596 Final Project

Kuang-Ming Chen, Mike Huang

Outline

- Introduction and Motivation
- Dataset
- Experiment Settings
- Evaluation
- Experiment Results
- Reference

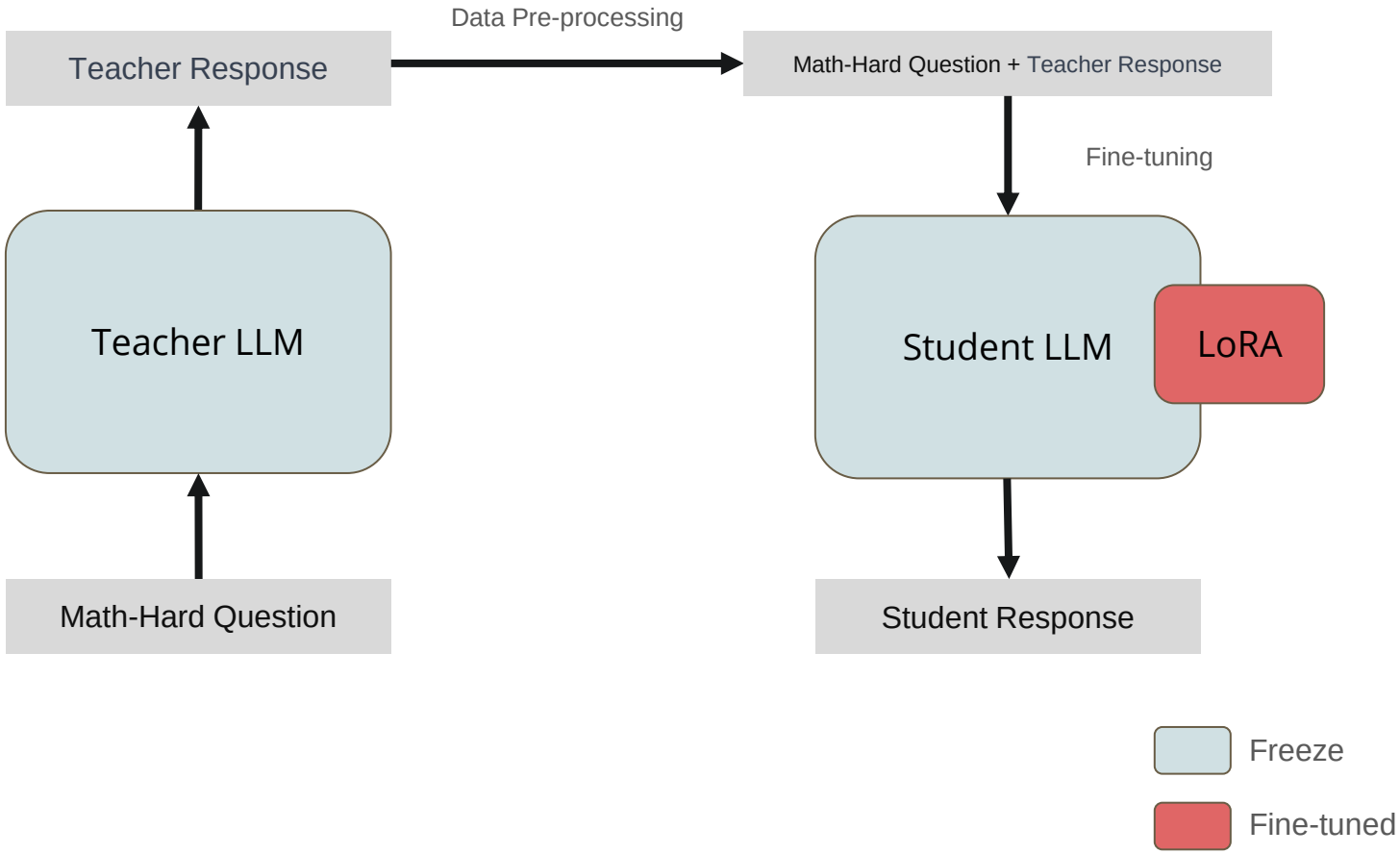
Introduction

- We are addressing a generation problem
- Goal: fine-tune LLM to solve difficult math questions by knowledge distillation

- Models: Llama-3.2-3B as teacher, Llama-3.2-1B as student
- Dataset: Math-Hard dataset
- Expectation: student model’s performance will fall between student model with direct fine-tuning and teacher model's inference results

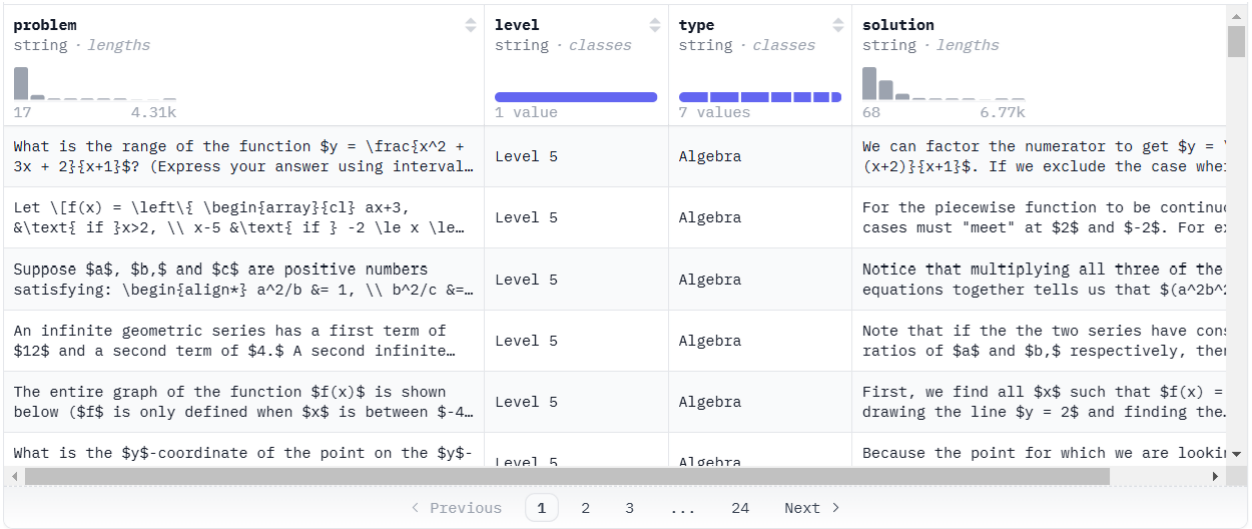
Motivation

- LLMs deploy on edge devices
- We need a smaller and high-performance model to solve real-world problems on these devices


Data Pre-processing
Teacher Response
Math-Hard Question + Teacher Response
Fine-tuning
Teacher LLM
Student LLM
LoRA
Math-Hard Question
Student Response
Freeze
Fine-tuned

Dataset

- Math-Hard dataset, 3.63k rows
- Use Problem and Solution columns in our project

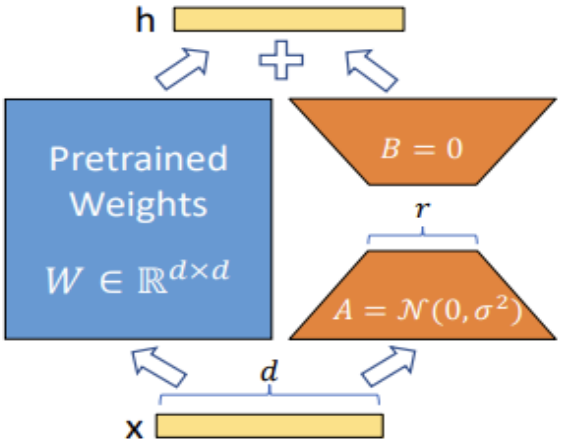


problem string · *lengths* 17 – 4.31k	level string · *classes* 1 value	type string · *classes* 7 values	solution string · *lengths* 68 – 6.77k
What is the range of the function $y = \frac{x^2 + 3x + 2}{x+1}$? (Express your answer using interval…	Level 5	Algebra	We can factor the numerator to get $y = (x+2)}{x+1}$. If we exclude the case whe
Let \[f(x) = \left\{ \begin{array}{cl} ax+3, &\text{ if }x>2, \\ x-5 &\text{ if } -2 \le x \le…	Level 5	Algebra	For the piecewise function to be continu cases must "meet" at 2 and -2. For e
Suppose a, $b,$ and c are positive numbers satisfying: \begin{align*} a^2/b &= 1, \\ b^2/c &=…	Level 5	Algebra	Notice that multiplying all three of the equations together tells us that $(a^2b^
An infinite geometric series has a first term of 12 and a second term of $4.$ A second infinite…	Level 5	Algebra	Note that if the the two series have con ratios of a and $b,$ respectively, the
The entire graph of the function $f(x)$ is shown below (f is only defined when x is between $-4…	Level 5	Algebra	First, we find all x such that $f(x) = drawing the line $y = 2$ and finding the.
What is the y-coordinate of the point on the y-	Level 5	Algebra	Because the point for which we are looki

‹ Previous 1 2 3 ... 24 Next ›

Experiment Settings

- Use LoRA module to reduce computing resources and time
- Fix LLM weight, only fine-tuning on LoRA parameters



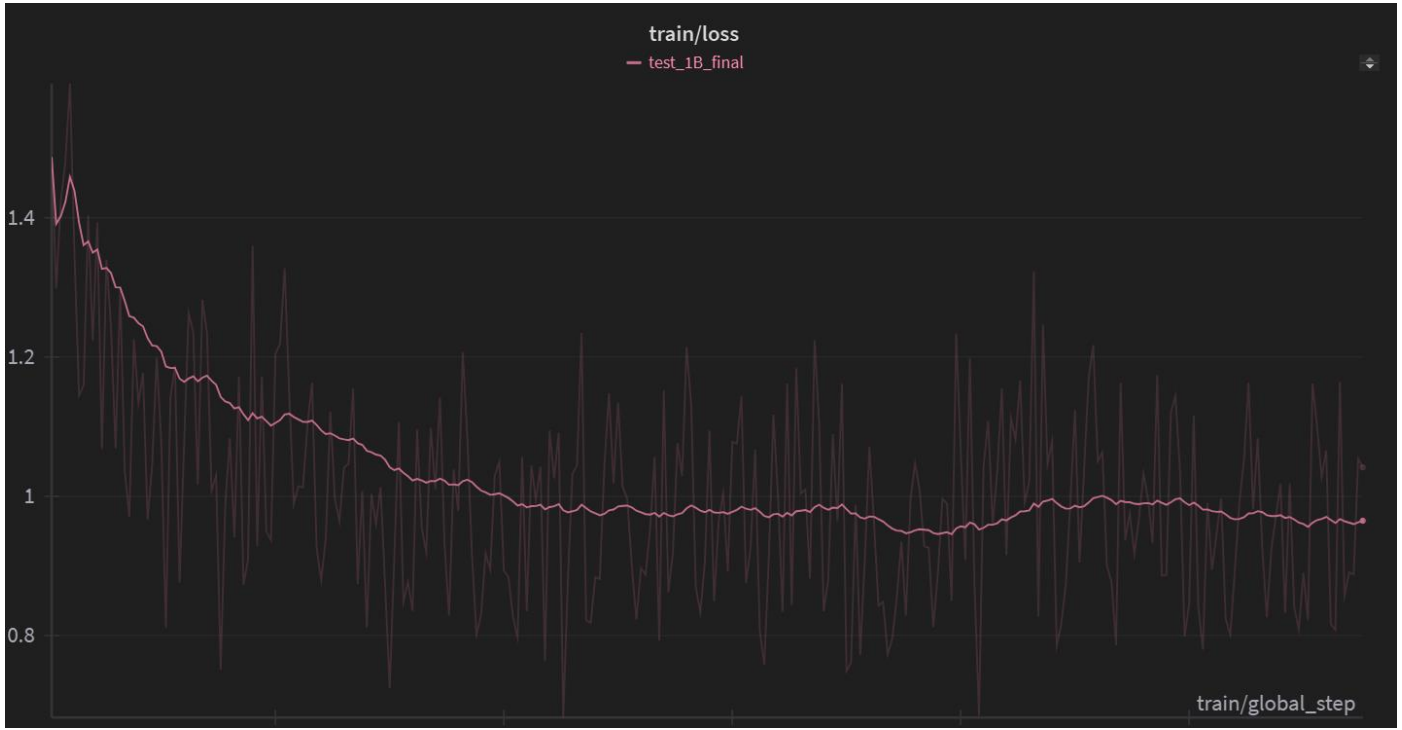
LoRA module

LLM: d1*d2 matrix
LoRA: d1*r + r*d2 matrix

Experiment Settings

- Batch_size = 2
- Epoch = 1
- Learning_rate = 2e-4
- Optimizer = AdamW
- Lr_scheduler = Linear
- LoRA_rank = 8

Experiment Settings



Training Loss (smoothing)

Evaluation

- MathQA dataset
- Multiple-choice questions, evaluate P(accuracy)

- Expectation: student model's performance falls between baseline 2 and 3

Baseline 4	3B fine-tune on Math-Hard and inference
Baseline 3	3B direct inference
Baseline 2	1B fine-tune on Math-Hard and inference
Baseline 1	1B direct inference

Experiment Results

Model	Fine-tune	Baseline	Accuracy on MathQA
Llama-3.2-3B-Instruct (LoRA)	Math_Hard	Baseline 4	0.3534
Llama-3.2-3B-Instruct	-	Baseline 3	0.3451
Llama-3.2-1B-Instruct (LoRA)	**Math_Hard_inf**	**Our Method**	**0.3333**
Llama-3.2-1B-Instruct (LoRA)	Math_Hard	Baseline 2	0.3303
Llama-3.2-1B-Instruct	-	Baseline 1	0.3283

Thank you!

Reference

https://huggingface.co/collections/meta-llama/llama-32-66f448ffc8c32f949b04c8cf

https://huggingface.co/datasets/lighteval/MATH-Hard

https://math-qa.github.io/

https://arxiv.org/pdf/2106.09685

https://github.com/EleutherAI/lm-evaluation-harness