自然语言处理
课程作业 PJ6

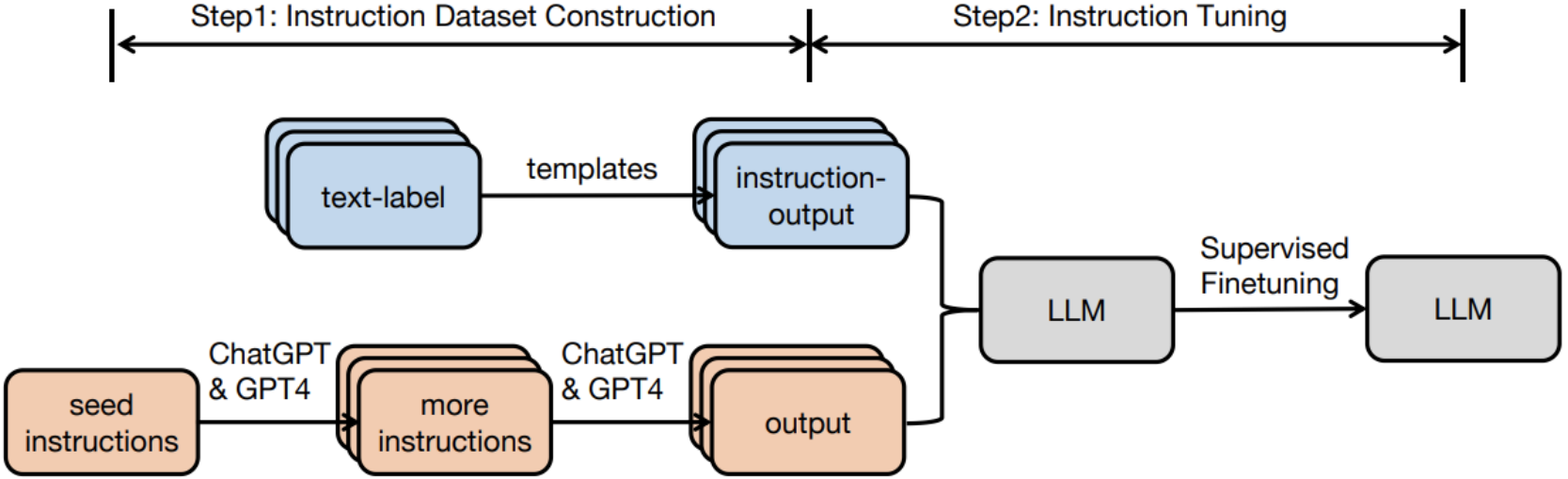
基于LLM微调的数学推理任务

助教：金森杰、仝竞奇、郭虹麟

What is **Instruction Tuning**?

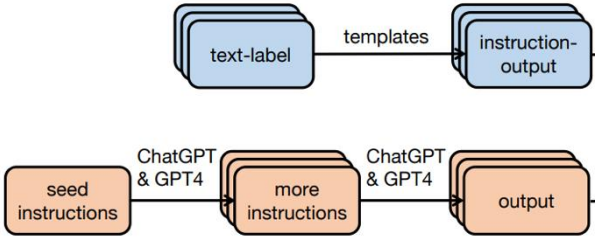
Which stage in large model training process?

General pipeline of **Instruction Tuning**

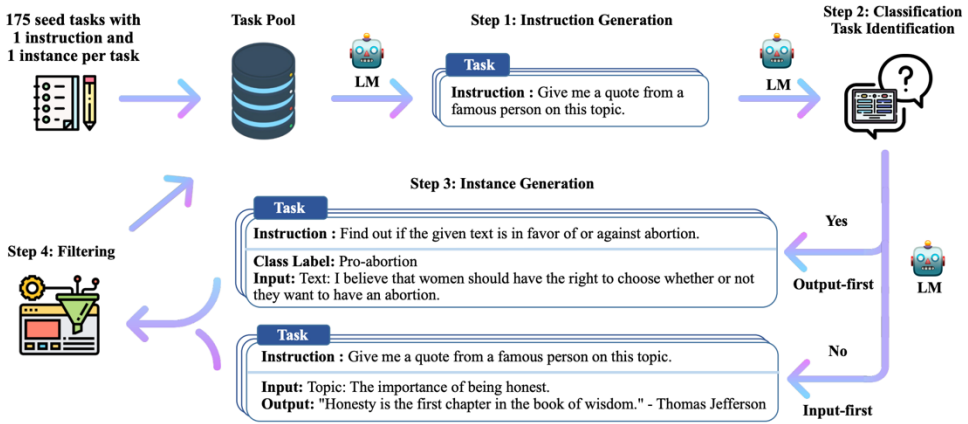


Instruction Tuning for Large Language Models: A Survey

Step1: Instruction Dataset Construction



- 指令微调数据构建
  - 开源/现有数据集根据模版构建
  - 人类标注
  - 大模型标注



Instruction: Given an address and city, come up with the zip code.
Input:
Address: 123 Main Street, City: San Francisco
Output: 94105

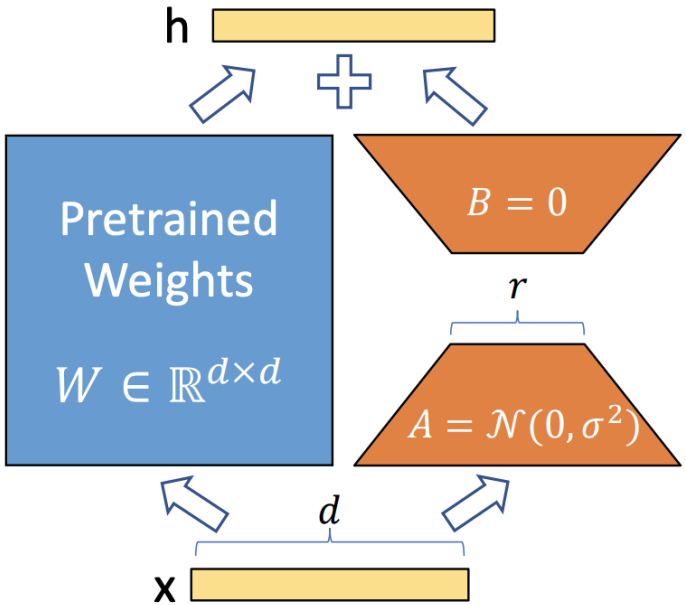
Instruction: I am looking for a job and I need to fill out an application form. Can you please help me complete it?
Input:
Application Form:
Name:_____ Age:_____ Sex:_____
Phone Number:_____ Email Address:_____
Education:_____ ...
Output:
Name: John Doe Age: 25 Sex: Male
Phone Number: ...

Instruction: How to write a code for converting degrees fahrenheit to celsius.
Input: Null
Output:

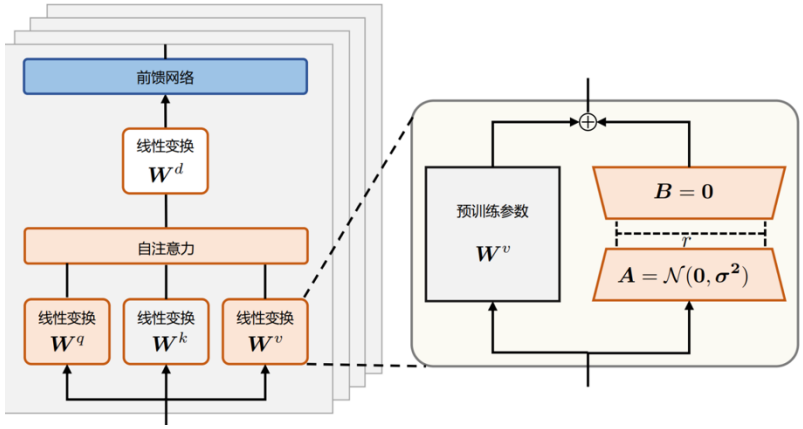
```
def convert_fahrenheit_to_celsius(fahr):
    celsius = (fahr - 32) * 5 / 9
    return celsius
```

Instruction Tuning for Large Language Models: A Survey
SELF-INSTRUCT: Aligning Language Models with Self-Generated Instructions

Step2: Instruction Tuning



$$h = W_0 x + \triangle W x = W_0 x + ABx$$

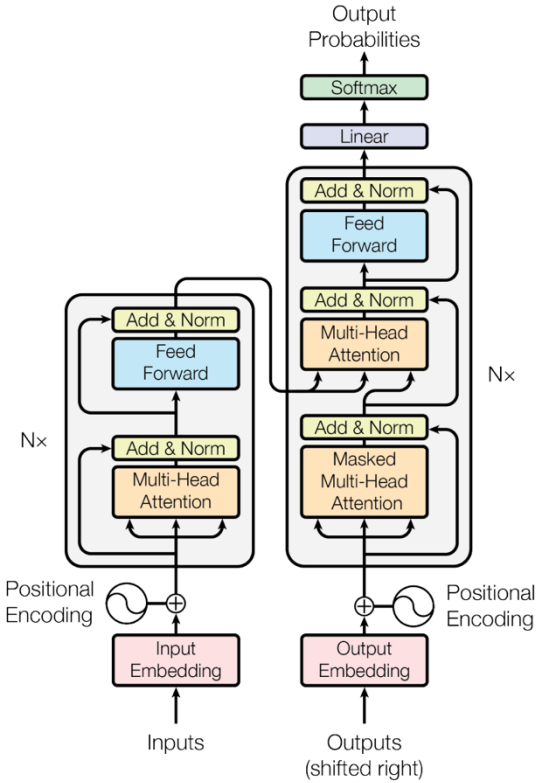


全参微调 VS 参数高效微调

Model&Method	# Trainable Parameters	WikiSQL Acc. (%)	MNLI-m Acc. (%)	SAMSum R1/R2/RL
GPT-3 (FT)	175,255.8M	**73.8**	89.5	52.0/28.0/44.5
GPT-3 (BitFit)	14.2M	71.3	91.0	51.3/27.4/43.5
GPT-3 (PreEmbed)	3.2M	63.1	88.6	48.3/24.2/40.5
GPT-3 (PreLayer)	20.2M	70.1	89.5	50.8/27.3/43.5
GPT-3 ($\text{Adapter}^{\text{H}}$)	7.1M	71.9	89.8	53.0/28.9/44.8
GPT-3 ($\text{Adapter}^{\text{H}}$)	40.1M	73.2	**91.5**	53.2/29.0/45.1
GPT-3 (LoRA)	4.7M	73.4	**91.7**	**53.8/29.8/45.9**
GPT-3 (LoRA)	37.7M	**74.0**	**91.6**	53.4/29.2/45.1

LORA: LOW-RANK ADAPTATION OF LARGE LANGUAGE MODELS

模型显存分析



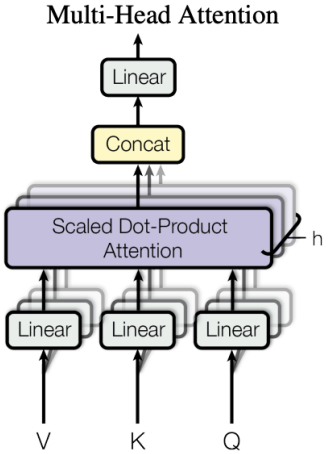
Attention Is All You Need

- 模型现存分析
- 环境上下文
- 静态显存
  - 模型参数消耗的显存，梯度消耗的显存，优化器消耗的显存；
- 中间激活状态消耗的显存
- 临时缓冲区占用的显存，以及其他零散的显存
  - 比较难以计算

假设模型的参数量为Φ。（以混合精度为例）

- 模型的参数消耗的显存：所有的模型参数需要存储到显存中，使用fp16存储，则需要消耗的显存为2Φ；
- 梯度消耗的显存：梯度和模型参数的量是完全相同的，使用fp16存储，则需要消耗的显存为2Φ；
- 优化器消耗的显存：以Adam为例，使用混合精度进行训练，Adam会备份一份fp32的模型参数，会以fp32存储averaged momentum和variance两部分，都和模型的参数量是相同的，即Φ。所以优化器消耗的显存为3∗4Φ。
- 这样，模型的参数消耗的显存、梯度消耗的显存、优化器消耗的显存分别为：2Φ、2Φ、4Φ∗3，总计为16Φ。以1.5B 的GPT-2为例，总大小为1.5B*(2 + 2 + 4*3) = 24G。

模型显存分析



首先面定义几个符号：

- b：表示batch_size；
- s：表示seq_length，为文本长度；
- h：表示hidden_dim，为隐藏层的维度；
- a：表示多头注意力中有多个头；
- h_a：表示hidden_dim_per_head，为多头注意力中每个头的隐藏层维度；

- 中间激活状态消耗的显存：是在前向传播的过程中，为了让后向传播完成计算，所需要保留的模型中间结果。

MHA 层需要保存的激活值，以及每个激活值的大小：

$$Q = x \cdot W_Q \quad : \quad 维度为\ [b, a, s, h_a] = [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{3}$$
$$K = x \cdot W_k \quad : \quad 维度为\ [b, a, s, h_a] = [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{4}$$
$$V = x \cdot W_v \quad : \quad 维度为\ [b, a, s, h_a] = [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{5}$$
$$Q \cdot K^T \quad : \quad 维度为\ [b, a, s, s], \quad 大小为\ 2bas^2\ 字节 \tag{6}$$
$$\text{softmax}(\frac{Q^T K}{\sqrt{d}}) \quad : \quad 维度为\ [b, a, s, s], \quad 大小为\ 2bas^2\ 字节 \tag{7}$$
$$\text{Dropout}\left[\text{softmax}(\frac{Q \cdot K^T}{\sqrt{d}})\right] \quad : \quad 维度为\ [b, a, s, s], \quad Dropout\ 层大小为\ bas^2\ 字节 \tag{8}$$
$$x_{self} = \text{Dropout}\left[\text{softmax}(\frac{Q \cdot K^T}{\sqrt{d}})\right] \cdot V \quad : \quad 维度为\ [b, a, s, h_a] = [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{9}$$
$$x_{self} \cdot W_o \quad : \quad 维度为\ [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{10}$$
$$\text{Dropout}(x_{self \cdot w_o}) \quad : \quad 维度为\ [b, s, h], \quad Dropout\ 层大小为\ bsh\ 字节 \tag{11}$$
$$x_{attn} = \text{LN}\left[\text{Dropout}(x_{self} \cdot w_o) + x\right] \quad : \quad 维度为\ [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{12}$$

FFN 层需要保存的激活值，以及每个激活值的大小：

$$x_{attn} \cdot W_{ff1} \quad : \quad 维度为\ [b, s, 4h], \quad 大小为\ 8bsh\ 字节 \tag{13}$$
$$\text{GeLU}(x_{attn} \cdot W_{ff1}) \quad : \quad 维度为\ [b, s, 4h], \quad 大小为\ 8bsh\ 字节 \tag{14}$$
$$x_{ffn} = \text{GeLU}(x_{attn} \cdot W_{ff1}) \cdot W_{ff2} \quad : \quad 维度为\ [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{15}$$
$$\text{Dropout}(x_{ffn}) \quad : \quad 维度为\ [b, s, h], \quad Dropout\ 层大小为\ bsh\ 字节 \tag{16}$$
$$\text{LN}\left[\text{Dropout}(x_{ffn}) + x_{attn}\right] \quad : \quad 维度为\ [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{17}$$

将 MHA 和 FFN 层全部加起来得到：

$$2bsh + 2bsh + 2bsh + 2bas^2 + 2bas^2 + bas^2 + 2bsh + 2bsh + bsh + 2bsh + 8bsh + 8bsh + 2bsh + bsh + 2bsh = 34bsh + 5bas^2$$

如果有 l 层 transformer，那么这 l 层 transformer 总的中间激活值占用的显存为：$l * (34bsh + 5bas^2)$

Attention Is All You Need

模型显存实例分析

MHA 层需要保存的激活值，以及每个激活值的大小：

$$Q = x \cdot W_Q \quad : \quad 维度为\ [b, a, s, h_a] = [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{3}$$
$$K = x \cdot W_k \quad : \quad 维度为\ [b, a, s, h_a] = [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{4}$$
$$V = x \cdot W_v \quad : \quad 维度为\ [b, a, s, h_a] = [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{5}$$
$$Q \cdot K^T \quad : \quad 维度为\ [b, a, s, s], \quad 大小为\ 2bas^2\ 字节 \tag{6}$$
$$\text{softmax}(\frac{Q^T K}{\sqrt{d}}) \quad : \quad 维度为\ [b, a, s, s], \quad 大小为\ 2bas^2\ 字节 \tag{7}$$
$$\text{Dropout}\left[\text{softmax}(\frac{Q \cdot K^T}{\sqrt{d}})\right] \quad : \quad 维度为\ [b, a, s, s], \quad Dropout\ 层大小为\ bas^2\ 字节 \tag{8}$$
$$x_{self} = \text{Dropout}\left[\text{softmax}(\frac{Q \cdot K^T}{\sqrt{d}})\right] \cdot V \quad : \quad 维度为\ [b, a, s, h_a] = [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{9}$$
$$x_{self} \cdot W_o \quad : \quad 维度为\ [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{10}$$
$$\text{Dropout}(x_{self \cdot w_o}) \quad : \quad 维度为\ [b, s, h], \quad Dropout\ 层大小为\ bsh\ 字节 \tag{11}$$
$$x_{attn} = \text{LN}\left[\text{Dropout}(x_{self} \cdot w_o) + x\right] \quad : \quad 维度为\ [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{12}$$

FFN 层需要保存的激活值，以及每个激活值的大小：

$$x_{attn} \cdot W_{ff1} \quad : \quad 维度为\ [b, s, 4h], \quad 大小为\ 8bsh\ 字节 \tag{13}$$
$$\text{GeLU}(x_{attn} \cdot W_{ff1}) \quad : \quad 维度为\ [b, s, 4h], \quad 大小为\ 8bsh\ 字节 \tag{14}$$
$$x_{ffn} = \text{GeLU}(x_{attn} \cdot W_{ff1}) \cdot W_{ff2} \quad : \quad 维度为\ [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{15}$$
$$\text{Dropout}(x_{ffn}) \quad : \quad 维度为\ [b, s, h], \quad Dropout\ 层大小为\ bsh\ 字节 \tag{16}$$
$$\text{LN}\left[\text{Dropout}(x_{ffn}) + x_{attn}\right] \quad : \quad 维度为\ [b, s, h], \quad 大小为\ 2bsh\ 字节 \tag{17}$$

将 MHA 和 FFN 层全部加起来得到：

$$2bsh + 2bsh + 2bsh + 2bas^2 + 2bas^2 + bas^2 + 2bsh + 2bsh + bsh + 2bsh + 8bsh + 8bsh + 2bsh + bsh + 2bsh = 34bsh + 5bas^2$$

如果有 l 层 transformer，那么这 l 层 transformer 总的中间激活值占用的显存为：$l * (34bsh + 5bas^2)$

- 中间激活状态消耗的显存：是在前向传播的过程中，为了让后向传播完成计算，所需要保留的模型中间结果。

计算的公式为 $l * (34bsh + 5bas^2) + 2bsh$，先来看这里的每个值分别是多少：

- l 为 tranformer 的层数，LLAMA-65B 为 80 层；
- b 为 batch_size，这里分别取 batch_size 为 16 和 1 计算两个不同 batch_size 下所需要显存的差别；
- s 为 seq_length，模型 LLAMA-65B 的最大长度为 2048；
- h 为隐藏层的维度，为 8192；
- a 为多头注意力层的 head 个数，为 64；

当 batch_size 为 1 时，计算结果如下：

$$\begin{aligned} &80 * (34 * 1 * 2048 * 8192 + 5 * 1 * 64 * 2048 * 2048) + 2 * 1 * 2048 * 8192 \\ =&80 * (570M + 1342M) + 33M \\ =&80 * 1912M + 33M \\ =&153G \end{aligned}$$

当 batch_size 为 16 时，计算结果如下：

$$\begin{aligned} &80 * (34 * 16 * 2048 * 8192 + 5 * 16 * 64 * 2048 * 2048) + 2 * 16 * 2048 * 8192 \\ =&80 * (9.13G + 21.47G) + 536M \\ =&80 * 30.6G + 536M \\ =&2448G \end{aligned}$$

Attention Is All You Need
https://mingchao.wang/4KTgtnFc/

数学推理任务

- 在过去几年里，大语言模型(LLM)在广泛的任务中取得了长足进展。最近，随着规模的扩大，LLM展现出了具备推理能力的潜力。
- 推理任务可能包括数学推理、逻辑推理、因果推理、视觉推理等

1. 经典问题 - 鸡兔同笼：
   笼子里有 23 只鸡和 12 只兔，问笼子里有多少个头和多少只脚?
2. GSM8K 问题：
   James buys 5 packs of beef that are 4 pounds each. The price of beef is $5.50 per pound. How much did he pay?

Meta-Question: James buys 5 packs of beef that are 4 pounds each. The price of beef is $5.50 per pound. How much did he pay?

Answer: He bought 5*4=20 pounds of beef. So he paid 20 * 5.5 = $110. The answer is: 110

- 指标

$$\text{Acc} = \textit{Corrections} / \text{Total}$$

数学数据集简介：

- **GSM8K数据集（Grade School Math）**是由 OpenAI 发布的小学数学题数据集。

Problem: Beth bakes 4, 2 dozen batches of cookies in a week. If these cookies are shared amongst 16 people equally, how many cookies does each person consume?
Solution: Beth bakes 4 2 dozen batches of cookies for a total of 4*2 = <<4*2=8>>8 dozen cookies
There are 12 cookies in a dozen and she makes 8 dozen cookies for a total of 12*8 = <<12*8=96>>96 cookies
She splits the 96 cookies equally amongst 16 people so they each eat 96/16 = <<96/16=6>>6 cookies
Final Answer: 6

Problem: Mrs. Lim milks her cows twice a day. Yesterday morning, she got 68 gallons of milk and in the evening, she got 82 gallons. This morning, she got 18 gallons fewer than she had yesterday morning. After selling some gallons of milk in the afternoon, Mrs. Lim has only 24 gallons left. How much was her revenue for the milk if each gallon costs $3.50?
Mrs. Lim got 68 gallons - 18 gallons = <<68-18=50>>50 gallons this morning.
So she was able to get a total of 68 gallons + 82 gallons + 50 gallons = <<68+82+50=200>>200 gallons.
She was able to sell 200 gallons - 24 gallons = <<200-24=176>>176 gallons.
Thus, her total revenue for the milk is $3.50/gallon x 176 gallons = $<<3.50*176=616>>616.
Final Answer: 616

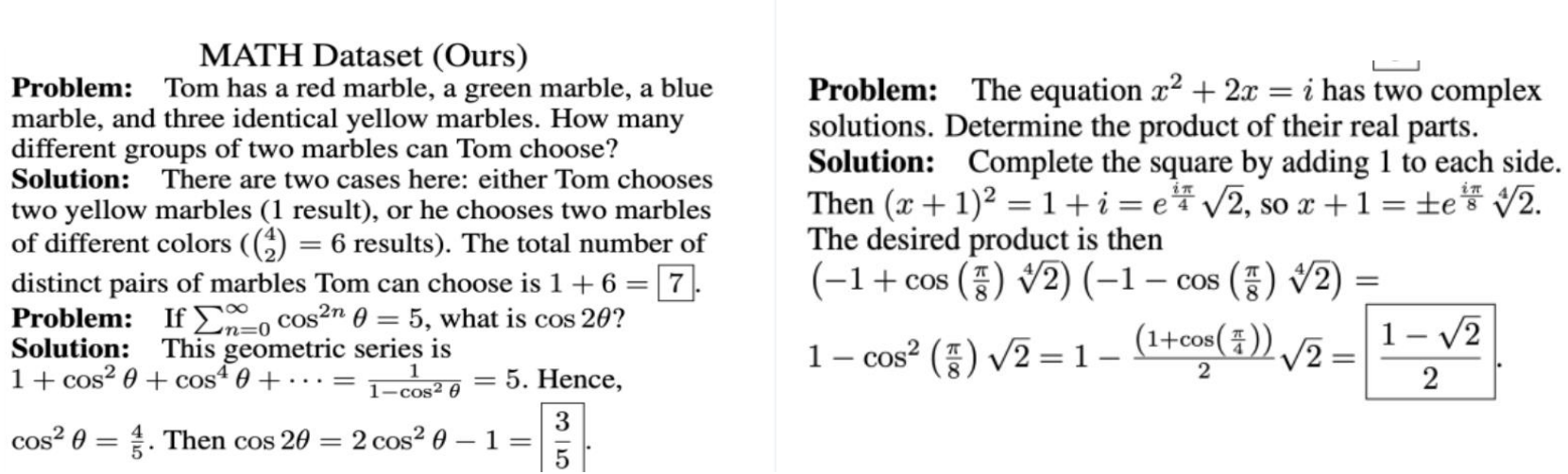
Problem: Tina buys 3 12-packs of soda for a party. Including Tina, 6 people are at the party. Half of the people at the party have 3 sodas each, 2 of the people have 4, and 1 person has 5. How many sodas are left over when the party is over?
Solution: Tina buys 3 12-packs of soda, for 3*12= <<3*12=36>>36 sodas
6 people attend the party, so half of them is 6/2= <<6/2=3>>3 people
Each of those people drinks 3 sodas, so they drink 3*3=<<3*3=9>>9 sodas
Two people drink 4 sodas, which means they drink 2*4=<<4*2=8>>8 sodas
With one person drinking 5, that brings the total drank to 5+9+8+3= <<5+9+8+3=25>>25 sodas
As Tina started off with 36 sodas, that means there are 36-25=<<36-25=11>>11 sodas left
Final Answer: 11

GSM8K由8.5K高质量的小学数学问题组成，这些问题都是由人类手写创造的。我们将这些问题分为7.5K训练问题和1K测试问题。这些问题需要2到8个步骤来解决，解决方法主要是使用基本的算术运算（+ - / *）进行一连串的基本计算，以得出最终答案。**一个聪明的中学生应该能够解决每个问题。**

https://raw.githubusercontent.com/openai/grade-school-math/refs/heads/master/grade_school_math

数学数据集简介：

- **MATH数据集**是一个由加州大学伯克利分校的研究团队开发的新数据集



MATH Dataset (Ours)

Problem: Tom has a red marble, a green marble, a blue marble, and three identical yellow marbles. How many different groups of two marbles can Tom choose?
Solution: There are two cases here: either Tom chooses two yellow marbles (1 result), or he chooses two marbles of different colors ($\binom{4}{2} = 6$ results). The total number of distinct pairs of marbles Tom can choose is $1 + 6 = \boxed{7}$.

Problem: If $\sum_{n=0}^{\infty} \cos^{2n}\theta = 5$, what is $\cos 2\theta$?
Solution: This geometric series is $1 + \cos^2\theta + \cos^4\theta + \cdots = \frac{1}{1-\cos^2\theta} = 5$. Hence, $\cos^2\theta = \frac{4}{5}$. Then $\cos 2\theta = 2\cos^2\theta - 1 = \boxed{\frac{3}{5}}$.

Problem: The equation $x^2 + 2x = i$ has two complex solutions. Determine the product of their real parts.
Solution: Complete the square by adding 1 to each side. Then $(x+1)^2 = 1 + i = e^{\frac{i\pi}{4}}\sqrt{2}$, so $x + 1 = \pm e^{\frac{i\pi}{8}}\sqrt[4]{2}$. The desired product is then $\left(-1 + \cos\left(\frac{\pi}{8}\right)\sqrt[4]{2}\right)\left(-1 - \cos\left(\frac{\pi}{8}\right)\sqrt[4]{2}\right) =$

$$1 - \cos^2\left(\frac{\pi}{8}\right)\sqrt{2} = 1 - \frac{\left(1+\cos\left(\frac{\pi}{4}\right)\right)}{2}\sqrt{2} = \boxed{\frac{1-\sqrt{2}}{2}}.$$

MATH数据集包含12,500个来自高中数学竞赛的挑战性问题，每个问题都有一个完整的逐步解决方案，这使得模型可以学习如何生成答案推导和解释。MATH数据集的问题覆盖了七个主要的数学领域，包括代数、几何、数论等，并且每个问题都标记了难度等级，从1到5，这允许对模型在不同难度和科目上的问题解决能力进行细致的评估。

https://github.com/hendrycks/math/

基于LLM微调的数学推理任务

- Prepare：
  - 统计算力；自行分组：根据算力，1-4人一组（最多4人）。
- 模型：
  - 中英文：
    - 1.Qwen-2.5 (0.5B、1.5B、3B、7B) https://huggingface.co/collections/Qwen/qwen25-66e81a666513e518adb90d9e
    - 2.InternLM2.5（1.8B，7B）https://huggingface.co/collections/internlm/internlm25-66853f32717072d17581bc13
  - 英文：
    - 1.Llama3.2 (1B、3B、3.1（8B)) https://huggingface.co/collections/meta-llama/llama-32-66f448ffc8c32f949b04c8cf
    - 2.Gemma2（2B，7B）https://huggingface.co/collections/google/gemma-2-release-667d6600fd5220e7b967f315
  - 数学增强：
    - 1.Deepseek-Math 7B https://huggingface.co/deepseek-ai/deepseek-math-7b-base
    - 2.Qwen2.5-Math （1.5B 7B）https://huggingface.co/collections/Qwen/qwen25-math-66eaa240a1b7d5ee65f1da3e
  - RL增强：
    - 1.Deepseek-Math-RL 7B https://huggingface.co/deepseek-ai/deepseek-math-7b-rl
  - **Qwen2.5-0.5B必须选项，2B及以下模型鼓励进行实验，算力有余力的可以尝试别的模型。**
- 任务：通过微调LLM，使得LLM在相应的数学测试集上获取高准确率。

基于LLM微调的数学推理任务

- 算法：（除了基本的SFT baseline，有复现的需标明引用、出处）
- 以下仅供参考
  - 数据增强方面：
    - 1.对数据集，输入输出的增强设计。
    - 2.扩充数据集（拒绝采样等方法，扩充原有数据集数目）。
  - 训练算法方面：
    - 1.lora高效训练
    - 2.离线/在线强化学习方法
  - Test time scaling方面：
    - 1.多数投票（majority voting）
    - 2.训练RM结合BoN和MCTS等

- 任务：通过微调LLM，使得LLM在相应的数学测试集上获取高准确率。

基于LLM微调的数学推理任务

- 评分指标:
  - 1.跑完Qwen2.5-0.5B CoT SFT，有测试结果，60分
  - 2.跑完自己方法，超过baseline，有实验过程依据，+20分
  - 3.其他看以下指标评判（剩余20分）：
    - 1.工作完整性;
    - 2.写作与表达;
    - 3.论文创新实践性;
- 提交文件：组号.zip（内含实验代码和报告，pdf格式）
  - 本次实验提交代码可不用ipynb格式
- Due:
  - 12月15日 **23:59**
- 迟交： 1天内分数*0.6，1天后为0分（如有任何特殊情况，请提前说明，无充分证明时不接受事后解释）

要求与说明：

- MATH数据集，测试集推荐使用MATH500子集：https://huggingface.co/datasets/qq8933/MATH500
- 1.主要关注Acc指标，相对Baseline要有提升；（Baseline是Qwen2.5-0.5B SFT）
- 2.以上只是推荐思路，可以自由发挥，合理即可，我们主要关注以下内容，不必卷模型/数据size和算力：
  - 1.鼓励方法创新性
  - 2.鼓励方法解决实际问题
  - 3.鼓励方法效果
  - 4.鼓励方法简单实用性（性价比）
- 3.学术诚信，实事求是
- 4.Nips模版:
  - 中文，正文不超过8页，要求严格按照学术论文格式（摘要、引言、相关工作、方法、实验、分析、结论、限制与讨论，引用和参考文献，每个人贡献（可不算在正文部分））；
  - 标题自拟；内容自定，与大模型推理（LLM Reasoning）相关即可；
  - 写作风格学术化，减少口语等现象；
  - 分析部分要求深入、有理有据；尽量避免简单记录实验现象；
  - 个人贡献，实事求是记录每个人的工作量；

联系我们

- 邮件:
  - 助教-金森杰：24110240038@m.fudan.edu.cn
  - 助教-仝竞奇：jqtong23@m.fudan.edu.cn
  - 助教-郭虹麟：hlguo24@m.fudan.edu.cn