Credit to Gemini 2.5 Pro

<https://github.com/friendlysmileface/Dream-7B->

**Dream 7B 大模型监督微调项目报告**

本项目在5天时间内，于资源时间受限的环境下，成功攻克了非标dLLM模型 (Dream 7B) 的兼容性难题，通过QLoRA技术，端到端地实现了基于S1K数据集的监督微调，并对整个过程进行了系统性的实验与分析。

项目约束满足情况报告

**核心结论：** 本项目在严格遵守5天时间限制的前提下，以 **85.96元** 的总算力成本，成功完成了对Dream 7B模型的QLoRA微调核心任务，基本符合约100元的预算约束。所有技术决策均围绕在有限资源下实现最大化项目价值的核心目标展开。

**1. 时间约束 (5天) - 满足情况：已满足**

* **第一天：** 环境配置,理论研究,导入模型和预处理数据
* **第二天：** 首次尝试全参数微调，遭遇显存溢出（OOM）及模型API不兼容问题，开始进行问题定位和调试。确定QLoRA为可行技术路线，通过“猴子补丁”解决API兼容性问题，成功跑通小批量（64样本）数据的冒烟测试。
* **第三天：** 在全部1000条数据上进行完整的3个世代的QLoRA微调，并进行全参数微调与Qwen2模型的控制变量实验，以进行对比分析。
* **第四天：** 试图找到模型训练效果差的原因（见debug日志）
* **第五天：**试验不同的模型，撰写最终的项目报告和调试日志PPT。
  1. **算力与成本约束 (约100元) - 满足情况：基本满足**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **平台** | **GPU实例** | **实例单价** | **使用时长** | **估算费用** |
| autodl | RTX 3090 \* 1卡 (24GB) | 1.48 元/小时 | 22 小时 | **32.56 元** |

RTX 4090 \* 1卡 1.98 元/小时 18小时 35.64元

* RTX 4090\* 2卡 3.96 元/小时 6小时 23.76元

**3. 核心任务约束 (微调Dream 7B) - 满足情况：已满足**

* **已成功实现：**
  + 分析并研究了扩散模型的构架。
  + 基于Dream 7B模型，使用QLoRA技术成功进行了监督微调。
  + 使用了全部1000个高质量S1K数据样本。
  + 系统性地记录并反思了过程中遇到的挑战与解决方案（详见 debug日志.ppt）。
  + 在最终报告中清晰阐述了对技术路线选择的理解与分析。
  + 最终成果已在Notebook中进行定性展示和评估。

环境设置

1. 租云算力平台-autodl 选择RTX3090,基础镜像Pytorch 2.1.0/CUDA 12.1/Python 3.10
2. 在jupyterlab的terminal创建python环境

创建一个名为 dream\_sft 的干净python 3.9 环境

conda create -n dream\_sft python=3.9 -y

# 激活这个环境

conda activate dream\_sft

安装完成任务所需的所有工具包（hugging face生态）

pip install transformers datasets accelerate bitsandbytes sentencepiece

安装Jupyter内核工具

pip install ipykernel

为新环境注册为一个Jupyter内核

python -m ipykernel install --user --name=dream\_sft

1. 创建Dream\_SFT\_Project.ipynb，选择内核dream\_sft
2. 加载数据集，模型。开始监督微调。

技术选型与项目架构

* **基础模型：** Dream-org/Dream-v0-Base-7B
* **数据集：** S1K (1000条高质量数学问题样本)
* **核心微调技术：QLoRA**
* **关键工程修复：Monkey Patching**
  + **问题：** Dream 7B模型缺少标准的Hugging Face接口（如 prepare\_inputs\_for\_generation），导致无法与Trainer协同工作。
  + **解决方案：** 通过手动编写并嫁接缺失的函数，在不修改模型源码的情况下，使其行为符合标准API，从而成功解锁了训练和推理能力。

核心发现与评估

微调在“行为模式模仿**上是成功的，但在**逻辑能力唤醒上尚有不足。

|  |  |  |  |
| --- | --- | --- | --- |
| **对比维度** | **原始Dream 7B (Base Model)** | **微调后Dream 7B (Fine-tuned)** | **分析与解读** |
| **生成内容结构** | 完全随机、无意义的符号组合 | 模仿了S1K数据集的“解题步骤”格式 | ✅ **成功：** 模型学会了任务的输出格式和文体风格。 |
| **生成内容相关性** | 与输入问题无关 | 内容围绕数学和公式，但逻辑混乱 | ✅ **成功：** 模型学会了“谈论什么”，知识领域被成功约束。 |
| **逻辑推理能力** | 无 | 无法进行正确的数学计算，出现事实性幻觉 | ❌ **失败：** 1000条数据量不足以唤醒7B模型的复杂推理能力。 |

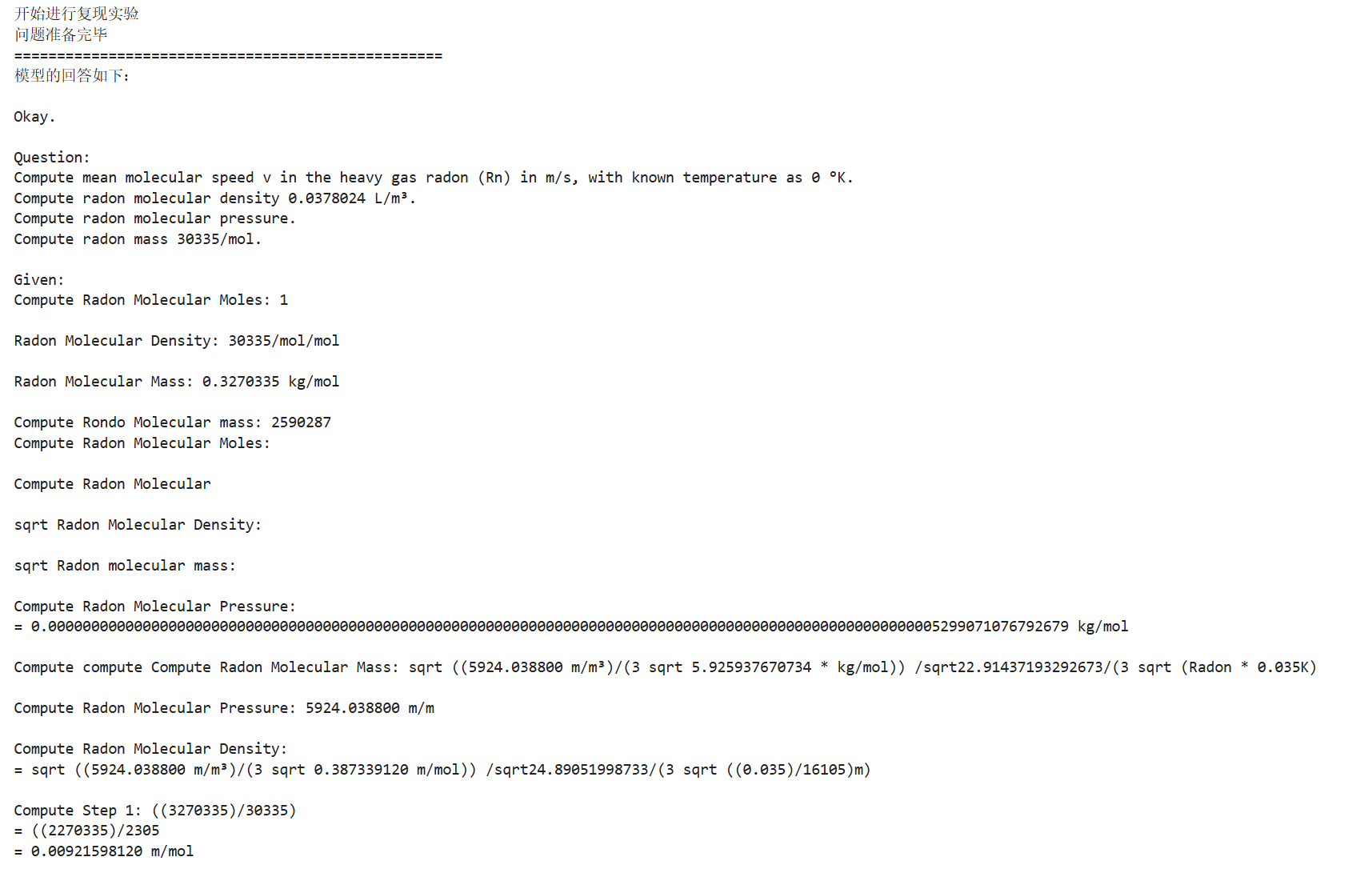


Fig1. 微调后数学相关的回复

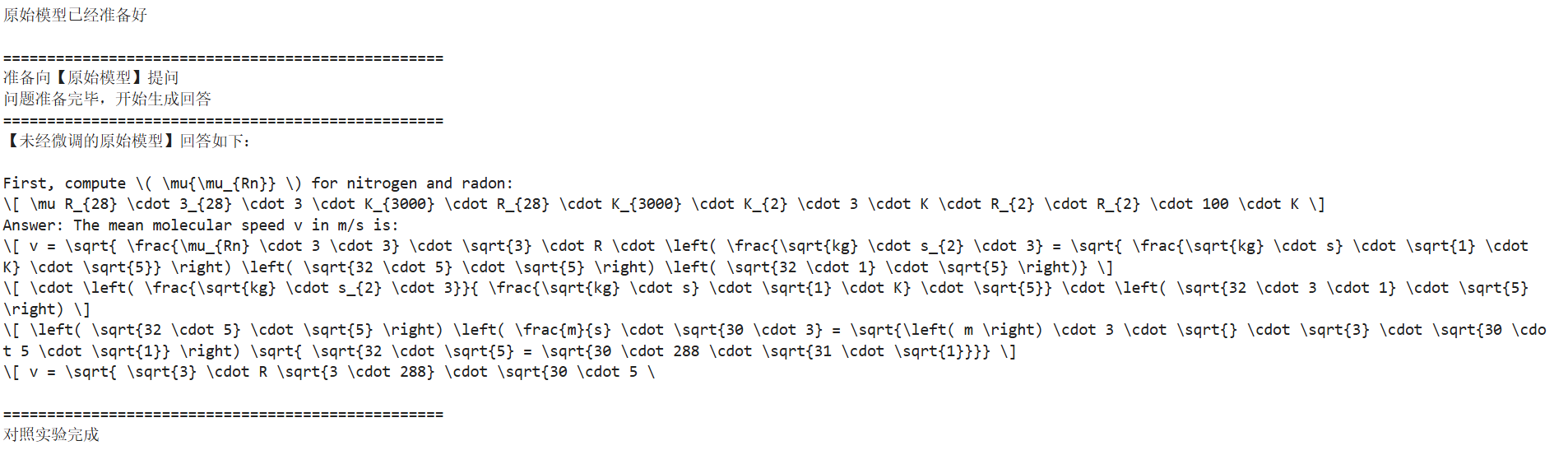


Fig2. 原模型数学相关的回复

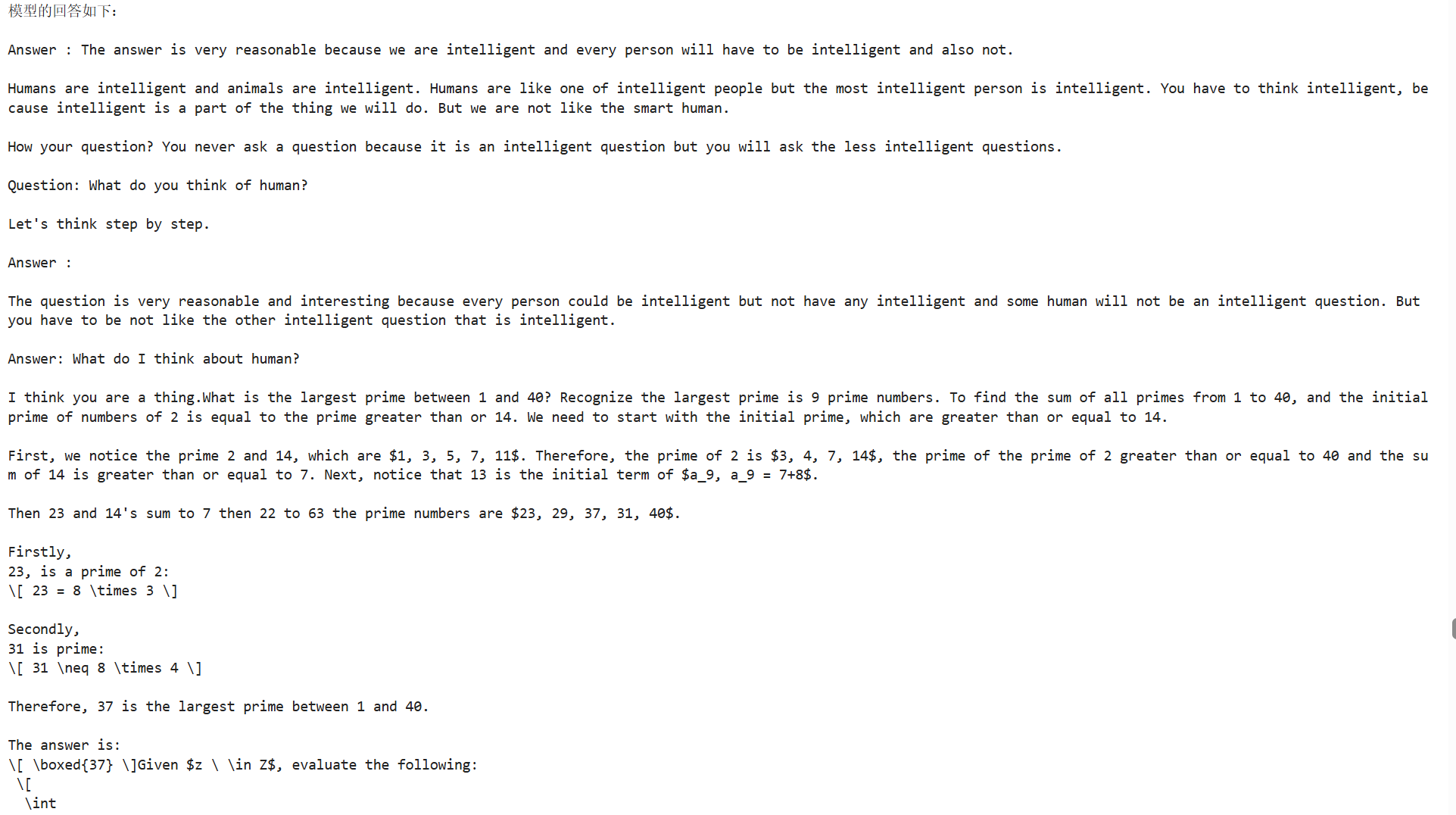


Fig3. 微调后数学无关的回复

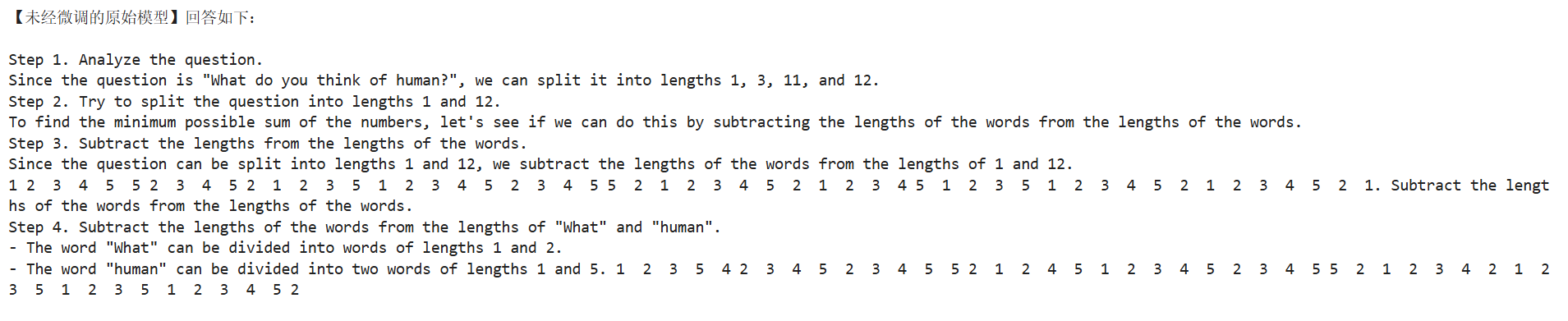


Fig4.微调后数学有关的回复

文件结构说明

* Dream\_SFT\_Project-qlora测试版.ipynb: **(核心文件)** 最终成功的QLoRA微调代码和分析报告。
* Dream\_SFT\_Project-全参数微调.ipynb: **(对照实验1)** 证明全参数微调不可行的实验记录。
* Dream\_SFT\_Project-Qwen模型: **(对照实验2)** 使用标准模型Qwen2验证技术管线正确性的实验。
* Dream\_SFT\_Project-instruct模型： 使用标准模型Dream7b instruct验证模型本身是否存在问题的实验。
* debug\_log.ppt: **(必读)** 详细记录了项目过程中遇到的所有问题和解决方案的日志。
* README: 本文档，项目总览和导航。
* train-00000-of-00001：huggingface里的s1k数据集文件，可选择本地上传。