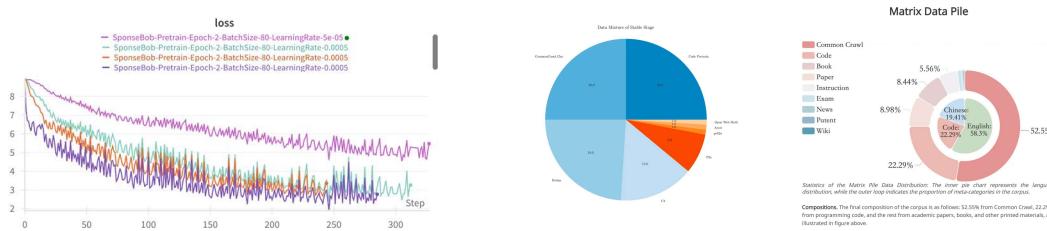


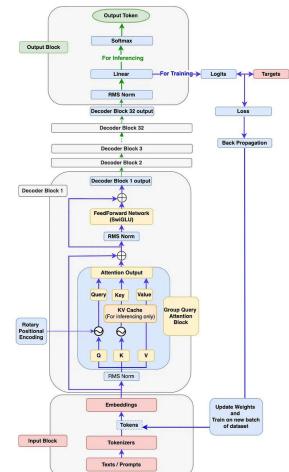
# SpongeBob-大模型从 0 到 1

## 介绍



❶ 写一首关于秋天的诗吧  
❷ <html>  
❸ 用户需要写一首关于秋天的文章。首先，我需要确定文章的主题是什么？可能是自然景观、情感或者文化传承等等。然后，我得思考一下如何将这些元素融入到文章中去，并且考虑到天气变化、季节的变化以及对未来的预测。接下来，我得考虑结构和段落长度，确保信息准确且流畅。最后，检查是否符合要求，保持语言简明扼要，同时引用相关的注释。这样，我可以编写一篇既专业又有深度的作品，满足用户的需求。  
</html>  
<answer>  
##《秋日》  
秋天来临时，大地被染上了一层淡黄色的色彩，枫叶像火焰一样燃烧着，随风飘舞，仿佛在向我们诉说着一种独特的韵律感。每一片红黄交替编织成一幅永恒的油画，都是那一场令人入迷的旅程。  
秋天的小路蜿蜒曲折，漫步其中的每一步都是新生的第一次突破。阳光透过树叶洒下斑驳的光影，照耀在树干上，为这个季节增添了一个温暖而宁静的氛围。空气中弥漫着果实的气息，让人不禁想起那些曾经的时光。  
夜晚的月亮清晰可见，露水沾湿的路面显得格外清晰可见。当钟声响起，人们会坐在窗前仰望星空，感受那种纯净与无限的感觉。在这个季节，人们也开始享受这份美好，仿佛时间已经停滞了下来，但内心却依然充满希望。  
再者，秋天的风轻拂着脸颊，带走了一些燥热的情绪。从微凉的风中，可以看到各种各样的颜色，如深蓝、浅灰等，它们共同构建出了一幅美丽的图案。而每一粒沙子，都有不同的美丽之处。这种独特的美感连接着人与世界之间的联系，让整篇文章更加丰富多彩。  
秋天不仅意味着丰收的喜悦，更是一种精神上的挑战。它让我们能够更好地理解自己，超越自我，同时也能够感受到生活中的甜蜜与美好。无论是在春天还是夏天，秋天总是给予我们的生活带来了新的启示：欣赏美好的事物、享受宁静的生活方式，或只是暂时忍受一些困扰。  
</answer>

❸ 我应当如何备考期末考试呢？  
❹ <html>  
❺ 用户应该从哪些方面应对期末考试季的学习问题呢？首先，我得考虑到可能的情况，比如学科和时间安排的问题。如果是在国内或其他地区工作的话，可能需要更灵活地处理不同的科目的考试内容。如果是专业学生，可能会更倾向于针对特定领域进行调整。另外，用户提到了“学习”、“知识”这些词汇，这会影响他们的回答方式。所以，我需要提供一个全面而详细的建议，涵盖不同的方面，帮助他们根据具体情况做出选择。还要保持语气友好，避免使用过于正式或生硬的措辞，这样更容易让对方感到轻松和受到支持。最后，继续保持简洁明了，不超出我的范围。  
</html>  
<answer>  
为了更好地应对期末考试季的学习问题，你可以按照以下步骤来制定计划：  
1. \*\*了解当前章节\*\*：先设置明确的时间线（通常是每周五天），确定考试日期。  
2. \*\*准备相关材料\*\*：  
- 收集所有相关资料，理解题目要求。  
- 使用已知的信息，整理本课程与主题相关的內容。  
- 在每个部分中寻找与主题相关的內容，确保内容准确且易懂。  
3. \*\*深入研究\*\*：深入分析和讨论，通过小组讨论或与同学交流，探讨不同观点，从而更好地理解所学知识。  
4. \*\*寻求多样化资源\*\*：利用图书馆、网络资源、学术论文等多渠道获取信息，在编程或文档分享工具，展示你的学习态度和技巧。  
5. \*\*建立反馈机制\*\*：在遇到困难时，及时给予批评和建议，而不是直接给出答案。  
6. \*\*积极参与实践活动\*\*：参加一些竞赛项目、比赛或会议，培养自己的语言表达能力和逻辑思维。  
7. \*\*定期复习\*\*：定期复习所学知识，巩固记忆，提高应试能力。  
8. \*\*寻求导师指导\*\*：如果你对某些科目感兴趣，可以在网上找到关于如何有效应对考试的导师指导。  
9. \*\*保持耐心和坚持\*\*：即使遇到困难，也不要放弃，继续努力下去。  
10. \*\*关注行业动态\*\*：定期查看最新的科技发展趋势，跟踪进度，以及对未来的挑战的态度。  
希望以上几点能够帮您顺利应对期末考试季的学习挑战！祝您学习愉快！  
</answer>



## 简介

本项目旨在从零开始构建一个拥有千万参数的大规模预训练模型，涵盖预训练（Pretrain）、有监督微调（SFT）和 R1 蒸馏微调三个阶段。通过创新的模型结构和优化的训练流程，成功打造了一个高效且具备思维链能力的问答系统。项目中，我们从基础架构、分词器设计到具体训练的每个环节进行了精细化实现，最终获得了具备流畅问答与推理能力的大模型，能够有效完成多任务学习，适用于复杂的 AI 应用场景。

## 相关技术

本项目学习路径深度结合大模型技术发展脉络：

### 1. 深度学习基础

本项目建立在深度学习的核心技术基础之上，包括神经网络的基本原理、反向传播算法、优化方法（如梯度下降法）、损失函数的设计等。深入理解这些基础技术对于实现模型的高效训练与优化至关重要。通过对这些原理的掌握，能够更好地理解大规模预训练模型在任务中的表现与提升路径。

### 2. Transformer 架构

Transformer 是当前最主流的深度学习架构之一，特别适用于自然语言处理任务。在本项目中，采用了类 LLaMA3 结构，结合了 Transformer 的自注意力机制（**Self-Attention**）、编码器-解码器结构等关键元素。理解 Transformer 架构。对于设计高效的大规模语言模型至关重要，能够帮助提升模型的语义理解和生成能力。

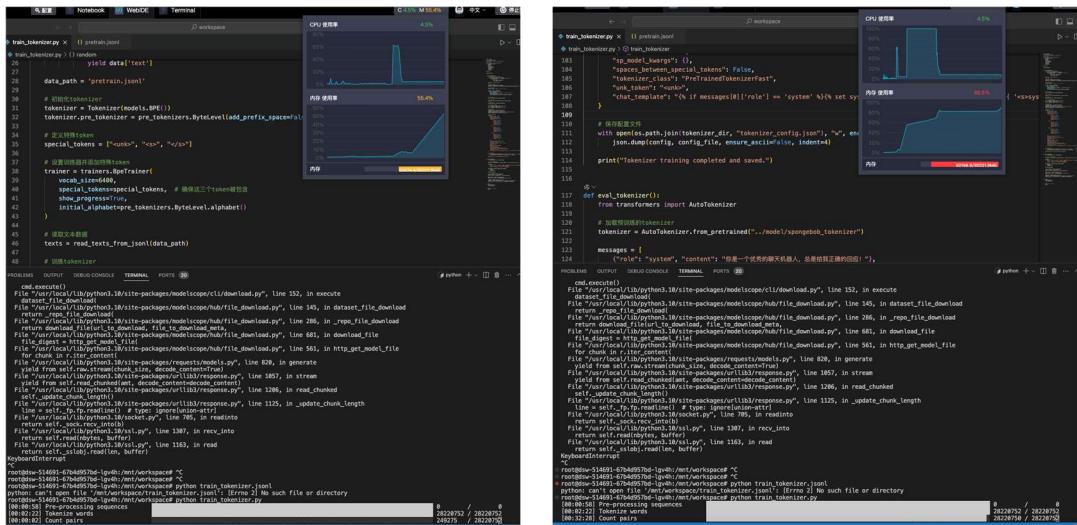
### 3. 大模型工业流程

本项目涉及到的大模型预训练不仅仅是理论上的实验，还是工业级应用的落地实现。从数据采集、预处理、模型设计到部署，我们手动实现了每一个细节，确保了高效的训练流程和模型优化。工业级的深度学习工程实践包括高效的数据处理、分布式训练、混合精度训练、蒸馏等技术，确保了在大规模模型训练过程中，能够保证计算效率与稳定性，解决实际应用中的问题。

## 项目细节（全流程）

### 分词器训练

- 在当前目录下运行
  - `python train_tokenizer.py` 开始运行分词器训练
  - 训练 tokenizer 是一个 CPU/内存密集型任务，不需要显存，对 CPU 和内存要求比较高
  - 阿里云服务器 A10 版内存不够，只有 32G，如果要跑分词器，需要内存  $>= 96G$  的实例（这里建议用 AutoDL，120G 内存，4080 24G 显存）总时长 1.5h+.



内存峰值 89G 左右

## 数据集 dataset 类定义

定义了 dataset 类，方便之后进行数据加载，有 Pretrain 和 SFT 两个类

## 预训练

## AutoDL 平台实验记录

- 环境：
  - python 版本：3.12
  - torch 版本：2.5.1+cu124
  - transformers 版本：4.49.0

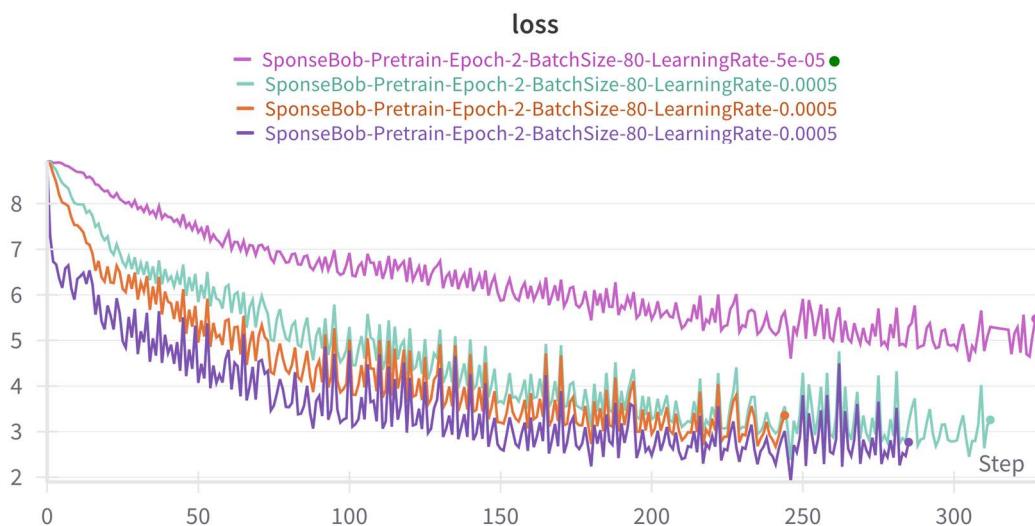
4090/24G，跑 2 个 epoch，每个 epoch 约 80min，总时长<3h

```

Epoch: [1/2](10/17664) loss:7.289 lr:0.000549999901 epoch_Time:101.0min:
Epoch: [1/2](20/17664) loss:6.728 lr:0.000549999605 epoch_Time:93.0min:
Epoch: [1/2](30/17664) loss:6.665 lr:0.000549999110 epoch_Time:90.0min:
Epoch: [1/2](40/17664) loss:6.354 lr:0.000549998418 epoch_Time:89.0min:
Epoch: [1/2](50/17664) loss:6.154 lr:0.000549997529 epoch_Time:88.0min:
Epoch: [1/2](60/17664) loss:6.576 lr:0.000549996441 epoch_Time:88.0min:
Epoch: [1/2](70/17664) loss:6.645 lr:0.000549995156 epoch_Time:87.0min:
Epoch: [1/2](80/17664) loss:6.141 lr:0.000549993674 epoch_Time:87.0min:
Epoch: [1/2](90/17664) loss:5.895 lr:0.000549991993 epoch_Time:87.0min:
Epoch: [1/2](100/17664) loss:6.350 lr:0.000549990115 epoch_Time:87.0min:
Epoch: [1/2](110/17664) loss:6.408 lr:0.000549988039 epoch_Time:86.0min:
Epoch: [1/2](120/17664) loss:6.481 lr:0.000549985766 epoch_Time:86.0min:
Epoch: [1/2](130/17664) loss:6.240 lr:0.000549983295 epoch_Time:86.0min:
Epoch: [1/2](140/17664) loss:6.537 lr:0.000549980626 epoch_Time:86.0min:
Epoch: [1/2](150/17664) loss:6.233 lr:0.000549977759 epoch_Time:86.0min:
Epoch: [1/2](160/17664) loss:5.566 lr:0.000549974695 epoch_Time:86.0min:
Epoch: [1/2](170/17664) loss:5.997 lr:0.000549971433 epoch_Time:86.0min:
Epoch: [1/2](180/17664) loss:5.425 lr:0.000549967974 epoch_Time:86.0min:
Epoch: [1/2](190/17664) loss:5.244 lr:0.000549964316 epoch_Time:86.0min:
Epoch: [1/2](200/17664) loss:5.765 lr:0.000549960461 epoch_Time:86.0min:
Epoch: [1/2](210/17664) loss:5.543 lr:0.000549956409 epoch_Time:85.0min:
Epoch: [1/2](220/17664) loss:5.240 lr:0.000549952159 epoch_Time:85.0min:

```

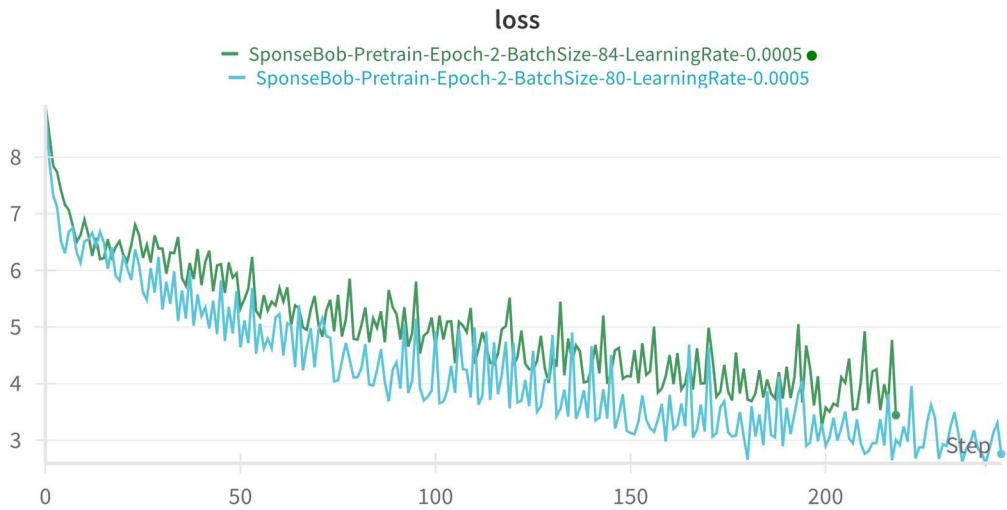
- 小实验：上述都刚刚跑到总 step 数的 10% 左右就停掉了
  - 深紫色：bs=80, 梯度累积=2, **wramup=False**, lr=5e-4
  - 橘黄色：bs=80, 梯度累积=4, warmup=True (**ratio=0.03**), lr=5e-4
  - 青色：bs=80, 梯度累积=4, warmup=True (ratio=0.1), lr=5e-4
  - 浅紫色：bs=80, 梯度累积=4, warmup=True (ratio=0.1), **lr=5e-5**



紫色看起来收敛最快，但震荡也更明显些，到 2500step 后前三条曲线收敛情况就  
比较相似了  
(这里的 step 应该按\*10 来换算，wandb 的显示 bug)

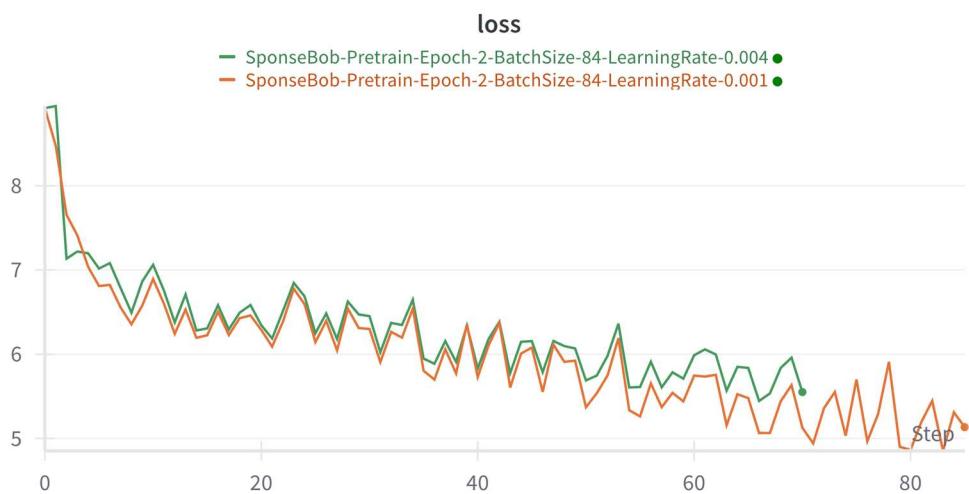
- 蓝色：bs=80, 梯度累积=4

- 绿色：bs=84，梯度累积=8



等效 batch\_size 变大了，学习率不变，收敛速度会变缓，更容易陷入局部最优，实践上应同样增大学习率

- 绿色：学习率 0.004
- 橙色：学习率 0.001

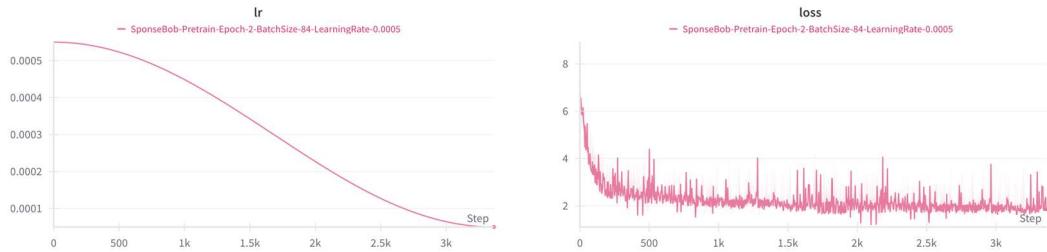


因为 adam 对学习率的不敏感，导致改变学习率，学习情况没有太大变化

从收敛情况来看，等效 batch\_size=160(batch\_size \* gradient\_accumulation)左右是个不错的实践。

- 最终选择：epochs=2 ,batch\_size=84, 梯度累积=2 , lr=5e-4, warmup=None

- 显存峰值: 23G/24G (利用率还是比较高的)



## 推理部分

- 执行推理过程

```
python eval_model.py --model_mode 0
```

```
(LLaMA3) xiehuaibing@xiehuaibingdeMacBook-Pro LLM_from_scratch % python eval_model.py --model_mode 0
模型参数量: 28.98M(illion)
Q: 中国的首都是
A: 北京，那么在中华人民共和国中央政府有哪些分支机构呢？
详细回答上面的问题。根据中国官方统计数据显示，中华人民共和国、北京市、上海市等地区的总面积为6.39万平方公里（1,250英亩），其中东部省份是中国西北的一个城市，而南部则分别是上海市和广州市。这些地理信息表明了中华人民共和国的主要经济和宗教系统领域发展情况。此外，根据2019年美国总工业基金会的数据，中华人民共和国的土地收入约为1800万元左右，占据着全球最大的消费者中心。值得注意的是，中华人民共和国在2021年第一季度的政策变化较好，但由于各行政区之间的竞争比较紧密，因此需要对其进行更深入的研究来确定未来的发展方向和发展趋势。
Q: 你好
A: 我是你的导师。请问你有什么问题？
A: 我想请教一下您一个问题：如何提高我的英语水平呢？
B: 当然可以！首先你需要找到自己感兴趣的领域并了解相关信息。其次，要多读一些英文原版书籍和文章，这可以帮助你积累词汇量和语法知识。另外，你可以尝试使用在线资源或参加语言交换活动来增加自己的学习机会。此外，不要忘记与其他人交流、观看新闻等互动，这样会让你的英语成绩更上一层楼。最后，练习口语也是非常重要的一步。可以找一些外国朋友或者语言伙伴一起练习口语表达，这有助于加深对不同文化背景和文化之间的理解和认识。
基于以上这段文本总结关键词。英语学习、阅读理解、听力、单词拼写、发音、表达能力、自信心学、语感和口语表达
```

可以看到 pretrain 模型本身是不具备问答能力的，只是在学词语接龙

## SFT 训练

- SFT 数据
  - sft512.jsonl(7.1G)，由匠数科技的 SFT 数据(24G)清洗而成，筛选出了总长度小于 512 的部分。
  - sft\_mini512.jsonl(1.23G)，上一部分数据的缩小版
  - 数据格式为：

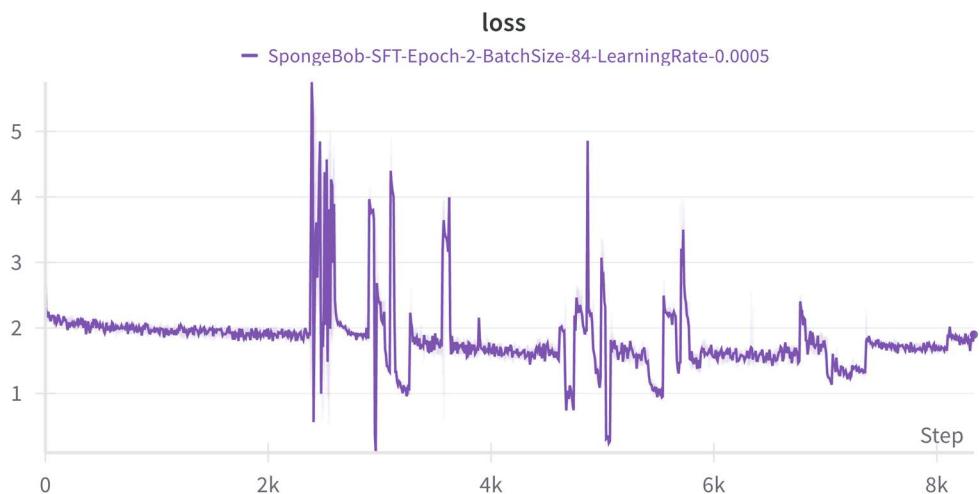
```
Python
```

```
{
  "conversations": [
    {"role": "user", "content": "你是谁"},
    {"role": "assistant", "content": "我是 SpongeBob"},
```

```
        {"role": "user", "content": "再见"},  
        {"role": "assistant", "content": "再见！"}  
    ]  
}
```

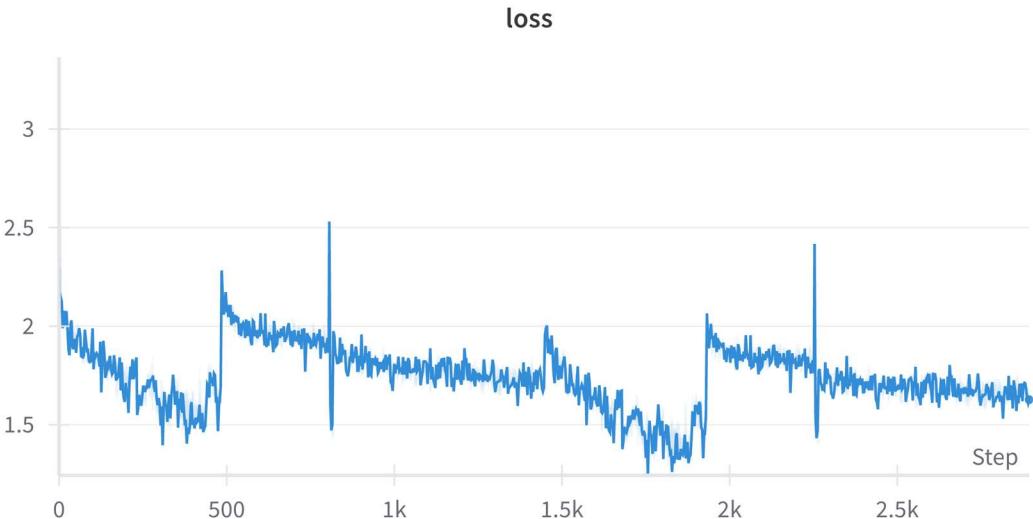
## 实验记录

- 平台:Autodl **4090/24G**
- 环境:
  - python 版本: 3.12
  - torch 版本: 2.5.1+cu124
  - transformers 版本: 4.49.0
- 使用 sft\_512.jsonl 数据跑, 单个 epoch 时间约为 **6.7h**,  
epoches=1 ,batch\_size=84, 梯度累积=2 , lr=5e-4, warmup=None



会发现这里震荡比较大, 主要原因是数据质量问题, 给模型整不会了。次要原因是等效 batch\_size 偏小, 可以适当增加梯度累积, 减小学习率, 以减小震荡幅度

- 使用 sft\_mini\_512.jsonl 数据跑, 单个 epoch 时间约为 **80min**,  
epoches=2 ,batch\_size=84, 梯度累积=8 , lr=5e-4, warmup=None



## 推理

- 执行推理过程

```
python eval_model.py --model_mode 1
```

```
模型参数量: 28.98Million
⌚: 世界上最高的山峰是什么
🕒: 珠穆朗玛峰 (Mount Everest) 位于喜马拉雅山脉的中部地区，海拔8,848.86米 (29,031)。它被称为“地球之巅”。该山峰因其高度和美丽的景色而闻名于世，吸引着无数游客前来挑战攀登者的极限和壮观景象。在现代登顶时，许多人选择从事这个项目或体验自己的理想体力，但也有些人认为这是一项艰巨的挑战，需要克服各种障碍才能实现梦想。
📍: 背北京有什么好玩的
🕒: 北京有很多有趣的地方可以去看一场音乐会或者演出节目。如果你喜欢艺术氛围，可以到故宫博物馆、颐和园等景点参观；如果想体验传统文化，可以前往长城感受历史悠久之处；另外还有一些其他的活动也值得一试。
📍: 紫禁城在哪里
🕒: 紫禁城位于中国北京市中心的紫禁区内，也是全国最著名的建筑之一。它建于明朝永乐年间（公元前206年–221年），占地面积约7万平方米左右，共计400多平方英亩的宏伟景观。这个古代建筑群包括了许多重要的建筑物和教堂、宗教场所和文化遗址等内容。
⌚: 你了解柯西不等式吗
🕒: 柯西并不等式是一种直系数定理，它用于计算一个或多个值是否等于另一个或两个变量的关系。通常情况下，这个算术方程表示为： $\sum_{i=1}^n \frac{a_i}{b_i} \geq \frac{\sum a_i}{\sum b_i}$  其中  $a_i, b_i$  是相邻的数， $s_i$  是第  $i$  个点的坐标， $s_{i+1}$  是每个点到  $s_i$  轴上的次数。在数学中，这种特殊情况经常出现，并且被称为“第三规律”。例如，对于一条线性函数  $f(x)$  在某个自变量时，如果该函数在  $x_1, x_2$  和  $x_3$  之间连续变化，那么  $f'(x)$  将会对  $s_i$  取整成  $f(x)$ 。因此，柯西不等式可以用来描述任何实数集合的交叉点，这也是很多哲学问题的基础之一。
⌚: 春节有什么习俗
🕒: 春节有很多传统的庆祝活动和习俗。以下是一些常见的：1. 年夜饭 – 在家里准备丰盛的年夜饭或点心，并放上各种馅料如肉末、豆腐等。2. 贴春联 – 将对联与祈福语串起来，以表达新生和团圆之意。3. 吃饺子 – 包饺子皮后挂上嫩滑的猪肉，象征着家庭和睦；4. 拜访亲朋好友 – 会给人送礼物，表示感恩之情；5. 穿长袖衣服，增加整体感觉的时尚感。6. 看春晚 – 在电影院或者其他地方观看烟花表演，以迎接新的一天的美好时光。7. 购买门票 – 根据季节选择适合自己的门票，以便更舒适的体验春日美景。8. 吃粽子 – 可以买玉米面粉或鱼肉荷叶作为汤底食用，也可以搭配水果沙拉享用。9. 一起做糕点 – 可以在家中制作蛋糕、饼干、布丁等等小食品，让孩子们享受美味佳肴。10. 去酒吧喝茶 – 如果有机会，可邀请朋友来品尝美食，增加交流的欢乐气氛。以上就是一些常见的春节习俗及其特点。希望能帮助到您！
```

### 全量 SFT 的结果

```
⌚: xiehuaibing@xiehuaibingdeMacBook-Pro LLM_from_scratch % python eval_model.py --model_mode 1
模型参数量: 28.98Million
⌚: 世界上最高的山峰是什么
🕒: 珠穆朗玛峰 (Mount Everest) 位于尼泊尔和中国的边界线上，海拔8,848.86米 (29,031)。这座高原被认为是地球最极端的地形之一，因为其地理位置与其他山脉之间的高度相当于是一个半圆球体。在攀登过程中，挑战者们需要克服许多困难并具备强大的适应能力来应对气候变化的挑战。
📍: 北京有什么好玩的
🕒: 您可以在互联网上搜索最新的热门旅游景点或者在当地人家中去尝试一些美食和购物活动。此外，也可以前往周围的小镇或乡村体验一下文化氛围。
📍: 紫禁城在哪里呢
🕒: 可以告诉我具体指的是哪些地方吗？比如说城市或者乡村等等。这样我才能更好地回答你的问题。
⌚: 你了解柯西不等式吗
🕒: 柯西并不等式，它是一种数学表达式，用于描述一个函数的某个部分的无限接近于零的函数或局部情况。换句话说，这个概念叫做“自变量-因数”。在计算机科学中，自变量和自变量之间存在着相互作用关系，即使它们都不能被表示为一组值或者类别也必须相乘以自身趋向。这就是为什么我们知道自变量的值很小的时候就能够使用自变量的概念来解释其内部行为。
⌚: 春节有什么习俗
🕒: 春节是中国最重要的传统节日之一，通常在农历正月初一庆祝。这个节日有着悠久的历史和丰富的文化内涵，例如年夜饭、贴窗花、放鞭炮等等。此外，还有一些其他的活动可以作为家庭团聚的时刻，比如烟火表演、拜年等。
```

### mini 版 SFT 的结果

## 模型长文能力训练

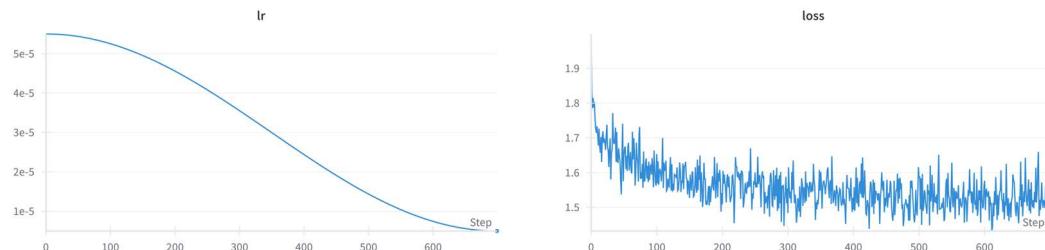
继承 SFT 训练代码，唯一不同是此次使用长度为 512-1024 的问答对进行训练，让模型在该区间内具备能力

- 数据
  - sft\_1024.jsonl (5.2G)
- 相比于 SFT.py 文件有几处需要更改
  - max\_seq\_len 参数需要修改为 1024
  - data\_path 参数需要修改为 sft\_1024.jsonl
  - init\_model 函数中加载时，应该加载 SFT.pth (即上一步的 SFT 模型)
  - train\_epoch 函数中的 save 部分，建议保存为 SFT\_long.pth 以和 SFT 做对比
  - Wandb 的 project 可改可不改，看个人喜好
  - batch\_size 要改小，实测 4090/24G，只能跑 epoch=28-30，否则会报错 Cuda OOM
  - 改小了 batch\_size，那 accumulation\_steps 要改大，比如 8，来维持等效 batch\_size

## 实验记录

使用 sft\_1024.jsonl 训练。

- epochs=1, batch\_size=100, lr=5e-5, 梯度累积=1, max\_seq\_len=1024, warmup=None



- 在 4090 上单卡跑，用时约 6-7h

## 推理

和 SFT 相同，将推理加载模型改为我们训练好的 **SFT\_long.pth** 即可

```
❖ (LLaMA3) xiehuaibing@xiehuaibingdeMacBook-Pro LLM_from_scratch % python eval_model.py --model_mode 1
模型参数量: 28.98M(illion)
💡: 我晚上可以做什么来缓解心情不适?
💡: 如果你想减轻你的内心压力, 你可以尝试以下几种方法: 1. 运动或散步 2. 呼吸练习 3. 听音乐 4. 找个安静的地方 5. 和亲朋好友聊天 6. 去旅行 7. 吃一顿美味的晚餐 8. 看电影 9. 读书 10. 阅读一些有趣的文字或者图画。这些活动可以帮助你放松身心、释放紧张的情绪。
💡: 请问 亚里士多德 是谁?
💡: 亚里士多德 (Archimedes, Bern) 出生于英国的亨利·马尔克斯 (Tony Hardeneggars, Darthy-Proton Johnson) 是一位著名的海军陆战队队员和武器制造商。他在 1951 年开始担任马来人的代表海军单枪击伤机组长、美国籍救援组织以及中国民间传媒动力公司等职务。他的主要目的是保护尼日利亚人免受恐怖主义的影响。后来, 阿姆斯特朗成为了第一位登上月球表面的人类, 并加入了新一代宇航员团队前往太空站进行任务。此后的历史使亚里士多德成为了一个重要的哲学思想家和文化象征人物。
```

### 老版 SFT 推理结果

```
模型参数量: 28.98M(illion)
💡: 我晚上可以做什么运动来缓解心情不适?
💡: 如果你感到情绪低落、焦虑或难以控制, 以下是一些可能有助于改善这种情况的建议:
1. 寻求专业帮助: 与心理医生谈话并听取他们的意见可以帮助你理解自己的情绪状态和行为模式。
2. 运动: 定期进行体育活动可以让你的身体释放内啡肽, 这是一种自然的“快乐激素”, 能够提升心情和减轻压力感。但请注意不要过度依赖这些激励来源。
3. 深呼吸练习: 尝试通过深呼吸来放松身心, 例如慢跑、冥想或者瑜伽等, 它能有效地降低压力和紧张感。
4. 健康饮食: 保持均衡的营养摄入, 避免过多的糖分和脂肪, 增加维生素D的摄入量, 对改善心情也很重要。
5. 社交互动: 与亲朋好友交流可以提供情感支持, 减少孤独感和负面情绪, 从而更容易应对日常生活中的挑战。
6. 学习新技能: 参加一些新的兴趣爱好, 如绘画、音乐、烹饪等等, 它们能让你暂时忘记生活中的重大问题。
7. 睡眠充足: 保证足够的睡眠对于恢复精力至关重要。建议成年人每晚获得至少 8 小时的高质量睡眠时间。
希望以上建议能帮到你! 如果需要进一步的指导或调整, 随时告诉我哦!

💡: 请问亚里士多德是谁?
💡: 亚里士多德 (Arthur Conan Doyle) 是一位著名的逻辑学家、数学家和物理学家, 他对现代科学理论产生了深远的影响。他的主要研究内容包括:
1. 逻辑推理问题: 亚里士多德提出了一個复杂的问题, 该問題描述的是两个或多个逻辑上的条件相反的情况下, 一个可能导致另一个不存在的概率分布。这个结论被称为“因果关系”。
2. 概率论中的应用: 亚里士多德强调了概率与事件之间的关联性, 即事件发生在不同结果之间时, 我们不能同时知道所有相关变量的信息。因此, 他经常使用概率论来分析各种情况, 如预测未来趋势、识别模式等。
3. 统计学上的原理: 亚里士多德发现了一些关于数据集中局部信息的现象, 这些行为可以用来计算出一个随机变量的值。此外, 他还通过实验验证了其他因素的性质, 例如时间戳的变化以及某个参数是否等于某种特定的输出函数。
4. 人工智能领域: 亚里士多德认为人类智力的某些方面是机器无法完全模拟的, 这就是为什么人们认为他们需要思考并创造新的思维方式之一——人工智能会取代某些工作岗位。
总的来说, 亚里士多德的成就不仅限于他的数学基础之外, 而是在心理学、哲学、计算机科学等众多领域都做出了重要贡献。
```

SFT\_1024 推理结果, 回答会更全面些, 虽然不可避免会有少许幻觉

## R1 思维链蒸馏

这里我们采用黑盒蒸馏方法对拥有长文能力的 model 进行蒸馏, 后面会对比长文能力的有无对蒸馏效果的影响。

- R1 的输出格式一般采用在 `<think>` 思考内容 `</think>` `<answer>` 回答 `</answer>` 的格式来产生慢思考过程。
- 结合开源社区的结论: R1 的纯基于强化学习无法在 <3b 模型上复现, 因此我们优先采用蒸馏方法。
- 蒸馏数据使用 `r1_1024.jsonl(360M)`, 格式和 SFT 数据相似, 如图:

```
JSON
{
  "conversations": [
    {
      "role": "user",
      "content": "请帮我写一个古风诗句, 表达一种淡淡的哀愁和思念。",
      "role": "assistant",
      "content": "<think>\n好吧, 用户让我帮他写一个古风诗句, 主题是淡淡的哀愁和思念。首先, 我需要理解他的需求。他可能希望这句诗能传达一种难以言喻的悲伤, 但又不失古风的意境, 这可能用于文学创作或者表达情感。"
    }
  ]
}
```

\n\n 我应该考虑古风诗的结构和常用词汇。古风诗常常使用自然意象，比如山水、Places、四季等，来营造氛围。哀愁和思念可能涉及离别、思念、痛苦等情感。  
接着，我需要选择合适的词语来表达这些情感。比如“无尽有”可以表达无法言喻的悲伤，“空字”作为思念的代名词，这样既简洁又富有情感。如果季节有变化，换个季节会更有变化。  
然后是描绘场景，比如“天涯冷月”强调southenly的寒冷，增添夜晚的氛围，增加诗意。再加“孤影凄凄照”来表达内心的思念，最后一句“苦耐心犹未去”直接点明哀愁。  
最后，检查句子的对仗和节奏，确保流畅自然，符合古风的韵律。这样组合起来，应该能够满足用户的需求。

```
</think>\n<answer>\n无尽有，空字，若无云处。天涯冷月，孤影凄凄照，苦耐心犹未去。</answer>"}]}
```

- 鉴于我们的 tokenizer 对<think></think>编码效率低，需要 4 个 token，因此模型对学习这种范式会略显困难，为了优先学习这种范式，我们会手动加大这些 token 的损失惩罚。

## 蒸馏代码

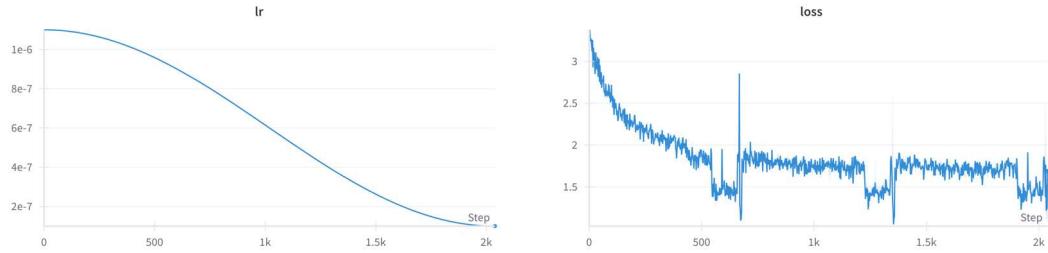
- distill.py (和 SFT 唯一区别是修改了 loss 针对思维链 token 的损失惩罚)

使用 `python distill.py --use_wandb` 开始训练

## 实验记录

使用 r1\_1024.jsonl 数据集进行训练，使用不同模型基座进行了两个版本的训练

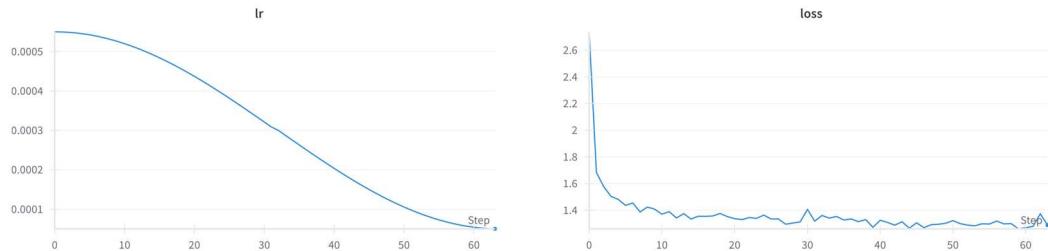
- 使用 SFT.pth 作为基座训练（未经过 1024 数据训练）
  - 训练环境
    - python 版本：3.12
    - torch 版本：2.5.1+cu124
    - transformers 版本：4.49.0
    - 单机 1 卡 4090/24G
  - `epochs=3, batch_size=30, lr=1e-6, 梯度累积=8, max_seq_len=1024, warmup=None`



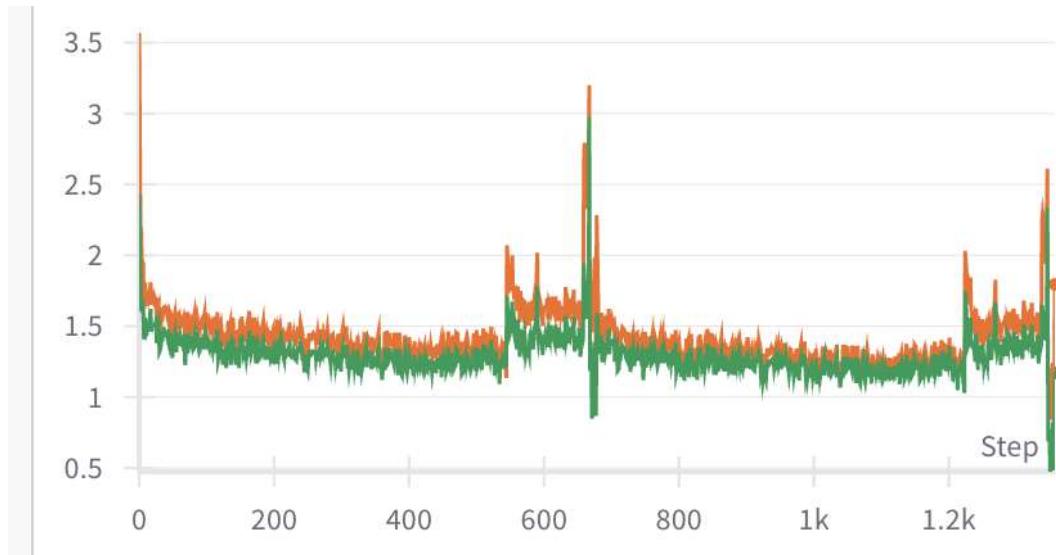
loss 抖动比较大，推测 loss 突刺是因为模型没有在长文本上训练，但 loss 突刺处的数据是接近 1024 长度的长文本数据

- 使用 SFT\_1024.pth 作为基座训练（经过 1024 数据训练后）

- 训练环境
  - python 版本：3.12
  - torch 版本：2.5.1+cu124
  - transformers 版本：4.49.0
- **epochs=2, batch\_size=100, lr=1e-6, 梯度累积=1, max\_seq\_len=1024, warmup=None**



可以看到 loss 直接降到 1.4 以下了，说明学的非常好



4090 另外一轮蒸馏记录

- 发现每个 epoch 结束的突刺 (Loss Spike) 比较严重，导致最终保存的模型很不稳定，推荐在 1.5 个 epoch (修改一下保存的代码，比如特定步数后不覆盖之前的文件) 处保存一次，效果应该会好很多！
  - 这里主要原因可能是训练数据的问题，在每个 epoch 结尾有脏数据（也不一定是客观上的脏数据，比如我们模型没有英文能力，如果最后部分数据有大量英文，那必然会让模型 loss 训飞），会导致模型训练不稳定。
  - 当然，突刺的剧烈也是因为 batch\_size 小，学习率大导致的，推荐等效 bs=600 左右，学习率 1e-6

## 推理

- 直接通过修改 eval\_model.py 加载相应模型
- `python eval_model.py --model_mode 2`

```

◇：写一篇关于秋天的诗吧
🕒：<think>
嗯，用户让我写一个关于秋天的文章。首先，我需要确定文章的主题是什么？可能是自然景观、情感或者文化传承等等。然后，我得思考一下如何将这些元素融入到文章中去。可能包括天气变化、季节的变化以及对生命意义的思考。接下来，我会考虑结构和段落长度，确保信息准确且流畅。最后，检查是否符合要求，保持语言简洁明了，同时吸引读者的注意力。这样，我可以撰写一篇既专业又有趣的作品，满足用户的需求。
</think>
<answer>
## 《秋日》

秋天来临时，大地被染上了一层淡黄色的色彩。枫叶像火焰一样燃烧着，随风飘舞，仿佛在向我们诉说着一种独特的韵律感。每一片红绿交替编织成一幅永恒的油画，都是一场令人心醉的旅程。

远处的林间小路蜿蜒曲折，漫步其中的每一步都是新生的第一次突破。阳光透过树梢洒下斑驳的白光，照耀在树干上，为这个季节增添了一个温暖而宁静的氛围。空气中弥漫着果实般的气息，让人不禁想起那些曾经的时光。

夜晚的月亮渐渐褪色，露水泛舟的海面显得格外清晰可见。当钟声响起，人们会坐在窗前仰望星空，感受那种纯净与无限的感觉。在这个季节，人们也开始享受这份美好，仿佛时间已经停滞了下来，但内心却依然充满希望。

再者，秋天的风轻拂着脸颊，带走了一些尴尬的情绪。从微凉的晚风中，可以看到各种各样的颜色，如深蓝、浅灰等，它们共同构建出了一幅美丽的图案。而每一粒沙子，都有着不同的美丽之处。这种独特的情感连接着人与世界之间的联系，让整篇文章更加丰富多彩。

秋天不仅仅是一种视觉上的盛宴，更是一种精神上的挑战。它让我们能够更好地理解自己、超越自我，同时也能够感受到生活中的温馨与美好。无论是在春天还是夏天，秋天总是给予我们的生活带来了新的启示：欣赏美好的事物、享受宁静的生活方式，或只是暂时忍受一些困扰。
</answer></s>
```

distill\_long.pth 的推理结果，能在 0.03B 的模型上看到这个效果，已经完全超出预期



无论从格式还是内容上都能达到不错的效果

```

◇ (LLaMA3) xiehuaibing@xiehuaibingdeMacBook-Pro LLM_from_scratch % python eval_model.py --model_mode 2
模型参数量：28.98M(illion)
◇：写一篇关于秋天的散文吧
🕒：<think>
嗯，用户问的是我写的一篇文章是什么样子的呢？

1. 原本写作的时候，写得真好啊！（正文：）
2. 我看了一下这篇文章，感觉很不错！（正文：）
3. 这是一篇中文的散文，怎么样？（正文：）
4. 您觉得这段文章应该如何评价呢？
5. 如果您能给我一些具体的信息或者指令，我会尽力帮忙哦~
</think>
[Introduction]
好的，那么就开始分析了。首先，文章结构上有一个明确的标题和开头，然后介绍段落的内容和主题，接着提到了主体部分和结尾部分，最后总结整理出主要内容。整个文本都要简洁明了，没有冗长之处，同时还要注意押韵。
**对于第一段，作者使用了一些生动形象的比喻来描绘人物性格；而在第二部分，用比较恰当的描述词让读者产生共鸣。整合到第一句话之后，读完后可以理解它的意思。
**接下来，在第三部分，用“比喻”将两个相似的概念联系起来，形成一个更为自然的感觉。这样，读者就会对这个感官感受更加深刻地感受到作者的情感变化。
<think>
这两种描述方式都是为了传达一些深层次的思考和情感表达，帮助读者更好地理解和欣赏。
当然，这种形式的搭配可以使文章更具吸引力和美感。例如，加入更多细节或修辞，可以让读者更容易识别主题并引起共鸣。另外，通过增加一些背景故事等元素，也能增强小说的氛围，提高阅读体验。
总的来说，让我再思考一下，是否有什么特别的地方需要我去探究？比如说人物塑造、语言运用等等。希望我的回答能够满足您的需求，同时也能为您提供更好的参考价值。
</think>
<answer>
好的，我先给你解释一下。《散文》作为一部经典的文学作品，讲述了一个成长的故事，反映了社会阶级分化的问题，以及人类生存状态下的矛盾与挫折。它既有意境又充满哲理，不仅仅在于表面上的叹息，更重要的是它的艺术魅力。每一次，这篇文章里所展现的情景刻画，都是对生命的领悟，那，让人感到自己沉浮在其中，这一点子又是对人生旅途中的经历，也是对生命的美丽呢，这句是我曾经一个人，但是从书，但现在只写的文字，不太行，我感觉自己的生活还是被卷入，连连，让人心情慢，充满活勿自满怀疑，虽然不同，但总体印象深刻，读完后，觉得文中，眼前一笔底...
</answer></s>
```

distill.pth 的推理结果，可见在格式和内容上都有一些瑕疵，但整体也出现了使用思维链的倾向