

minimind训练大模型过程

目录

- **0. 下载数据集**
 - 0.1. 构建GPU的pytorch镜像（需要有Nvidia的Docker版本以及基础设施）
 - 0.2. 下载数据集
 - 0.3 数据集的样本长什么样子
- **预训练（Pretain）**
 - 1.1 文本如何变成embedding的索引数组和embedding张量
 - 在下一步就要进行分词时数据集的样子：
 - 在转变为embedding的索引数组后：
 - 在转变为embedding的索引数组且padding后：
 - 索引数组还伴随一个loss掩码数组：
 - 构造输入与监督输出：
 - 1.2 执行单机四卡训练
 - 1.3 评估模型
- **单轮对话的全参数监督微调（Full Supervised Fine-Tuning, Full-SFT）**
 - 2.1 数据集如何变成embedding的索引数组和embedding数组
 - 数据集sft_data_single.csv中的一条数据被转换为了预训练中类似的一条普通的样本：
 - minimind的chat_template:
 - 2.2 执行单机四卡单轮对话全参数监督微调
 - 2.3 评估模型的单轮对话
- **多轮对话的全参数监督微调**
 - 3.1 数据集如何变成embedding的索引数组和embedding数组
 - 3.3 执行单机四卡多轮对话全参数监督微调
- **DPO训练**
 - 4.1 DPO损失函数
 - 4.2 一步DPO训练
 - 4.3 执行DPO训练

minimind地址：<https://github.com/jingyaogong/minimind>

使用的commit：6e9cd28ef9b34a0a10afbfd6f59e65cb6e628efb

0. 下载数据集

0.1. 构建GPU的pytorch镜像（需要有Nvidia的Docker版本以及基础设施）

构建镜像的Dockerfile：（要删除项目中requirements.txt自带的numpy和torch行，基础镜像里已经有了）

```
</> Docker
1 FROM nvcv.io/nvidia/pytorch:24.01-py3
2
3 WORKDIR /minimind
4
5 COPY requirements.txt .
6 RUN pip3 install -i https://mirrors.tuna.tsinghua.edu.cn/pypi/web/simple --no-cache-dir -r requirements.txt
7
8 ENTRYPOINT ["/bin/bash"]
```

启动镜像：

```
</> Bash
1 docker run --network=host --ipc=host --ulimit memlock=-1 --ulimit stack=67108864 --gpus '"device=0,1,2,3"' --rm -d -v ./minimind -w /minimind --name mimmind_dev minimind:0.0.1 sleep infinity
2 docker exec -it mimmind_dev /bin/bash
```

0.2. 下载数据集

安装git-lfs，然后执行：

```
</> Bash
1 git clone https://hf-mirror.com/datasets/jingyaogong/minimind_dataset
```

使用其他下载方式很麻烦且很慢，例如百度网盘

0.3 数据集的样本长什么样子

分词器的数据集：tokenizer_train.jsonl

```
</> Bash
1 {"text":
  "\u597d\u7684\u3002\u73b0\u5728\u8bf7\u4f60\u5c06\u8fd9\u4e2a\u6587\u672c\u4e2d\u7684\u6240\u6709\u7684\u9017\u53f7\u90fd\u66ff\u6362\u6210\u7a7a\u683c\u3002
  \u597d\u7684\u201c\u8bf7\u7a0d\u7b49\u4e00\u4e0b\u201d\u73b0\u5728\u6211\u4f1a\u5c
```

06\u6587\u672c\u4e2d\u7684\u6240\u6709\u9017\u53f7\u66ff\u6362\u4e3a\u7a7a\u683c\u3002\u5904\u7406\u540e\u6587\u672c\u4e3a\u55ef1a\u8fd9\u662f\u4e00\u4e2a\u53e5\u5b50

\u76ee\u7684\u662f\u770b\u770b\u662f\u5426\u53ef\u4ee5\u6b63\u786e\u5730\u4ece\u8fd9\u4e2a\u53e5\u5b50\u4e2d\u5220\u9664\u5173\u952e\u8bcd\u3002\u3002\u5904\u7406\u7ed3\u679c\u5982\u4f55\u55ef1f"}

2 {"text":

"\u5e2e\u6211\u56de\u7b54\u4e00\u9053\u5386\u53f2\u9898\u76ee\u3002\u6e05\u671d\u65f6\u671f\u7684\u516b\u65d7\u5171\u6709\u591a\u5c11\u65d7\u4eba\u55ef1f

\u6e05\u671d\u65f6\u671f\u516b\u65d7\u65d7\u4eba\u603b\u6570\u7ea6\u4e3a200\u4e07\u4eba\u5de6\u53f3\u55ef0c\u5176\u4e2d\u6b63\u9ec4\u65d7\u3001\u6b63\u84dd\u65d7\u3001\u6b63\u767d\u65d7\u3001\u6b63\u7ea2\u65d7\u7684\u4eba\u6570\u8f83\u591a\u55ef0c\u5176\u4ed6\u65d7\u7684\u4eba\u6570\u8f83\u5c11\u3002"}

3 {"text":

"\u55ef\u55ef\u8c22\u8c22\u4f60\u4ecb\u7ecd\u7684\u505a\u6cd5\u5f88\u8be6\u7ec6\u55ef0c\u4f46\u6211\u4e0d\u559c\u6b22\u5403\u9e21\u86cb\u55ef0c\u6709\u6ca1\u6709\u5176\u4ed6\u83dc\u505a\u6cd5\u80fd\u4ecb\u7ecd\u4e00\u4e0b\u55ef1f

\u5f53\u7136\u55ef0c\u4f60\u53ef\u4ee5\u8bd5\u8bd5\u9178\u8fa3\u571f\u8c46\u4e1d\u8fd9\u9053\u83dc\u3002\u6750\u6599\u55ef1a\u571f\u8c46\u4e2a\u3001\u7ea2\u6912\u4e2a\u3001\u9752\u6912\u4e2a\u3001\u5927\u8471\u6839\u3001\u918b\u3001\u751f\u62bd\u3001\u76d0\u3001\u9e21\u7cbe\u3001\u6599\u9152\u505a\u6cd5\u55ef1a\u6750\u571f\u8c46\u53bb\u76ae\u55ef0c\u5207\u6210\u4e1d\u55ef1b\u7ea2\u6912\u548c\u9752\u6912\u5207\u6210\u7ec6\u4e1d\u55ef1b\u5927\u8471\u5207\u6bb5\u5907\u7528\u3002\u6700\u9505\u51c9\u6cb9\u55ef0c\u6cb9\u70ed\u540e\u653e\u5165\u571f\u8c46\u4e1d\u55ef0c\u7178\u7092\u81f3\u53d8\u8f6f\u3002\u6750\u5165\u7ea2\u6912\u3001\u9752\u6912\u548c\u5927\u8471\u6bb5\u55ef0c\u7ee7\u7eed\u7178\u7092\u81f3\u719f\u3002\u6750\u52a0\u5165\u9002\u91cf\u7684\u76d0\u3001\u9e21\u7cbe\u3001\u6599\u9152\u548c\u751f\u62bd\u55ef0c\u7ffb\u7092\u5747\u5300\u3002\u6700\u540e\u55ef0c\u52a0\u5165\u9002\u91cf\u7684\u918b\u55ef0c\u7ffb\u7092\u5747\u5300\u5373\u53ef\u3002\u6750\u5c0f\u8d34\u58eb\u55ef1a\u6750\u571f\u8c46\u5207\u4e1d\u65f6\u55ef0c\u53ef\u4ee5\u653e\u5165\u6de1\u76d0\u6c34\u4e2d\u6ce1\u4e00\u4e0b\u55ef0c\u8fd9\u6837\u53ef\u4ee5\u53bb\u9664\u591a\u4f59\u7684\u6dc0\u7c89\u3002\u6750\u716e\u571f\u8c46\u4e1d\u65f6\u55ef0c\u4e00\u8981\u716e\u5f97\u592a\u4e45\u55ef0c\u4ee5\u514d\u70c2\u7cca\u3002\u6750\u52a0\u5165\u918b\u7684\u65f6\u5019\u55ef0c\u6839\u636e\u81ea\u5df1\u7684\u53e3\u5473\u591a\u5c11\u6765\u8fdb\u884c\u8c03\u8282\u55ef0c\u4e00\u822c\u6765\u8bf4\u55ef0c\u76d0\u4e0e\u918b\u7684\u6bd4\u4f8b\u4e3a1:1\u3002\u6750\u5982\u679c\u559c\u6b22\u8fa3\u5473\u53ef\u4ee5\u52a0\u5165\u4e00\u4e9b\u5e72\u8fa3\u6912\u4e1d\u3002\u6750\u5e0c\u671b\u4f60\u4f1a\u559c\u6b22\u8fd9\u9053\u9178\u8fa3\u571f\u8c46\u4e1d\u55ef01"}

4 {"text": "How can AI technology be used to address climate change? Climate change is one of the most significant challenges facing our planet today, and it is increasingly clear that we need to take bold action to address it. Artificial Intelligence (AI) technology has the potential to be a powerful tool in the fight against climate change. Here are some ways that AI technology can be used to address climate change:\n\n1. Energy optimization: AI can be used to optimize energy consumption by analyzing data from sensors and other sources to identify

areas of waste and inefficiency. This can help reduce energy consumption and lower greenhouse gas emissions.

2. Smart grids: AI can be used to manage smart grids, which can help optimize energy distribution, reduce energy waste, and enable the integration of renewable energy sources into the grid.
3. Renewable energy optimization: AI can be used to optimize the use of renewable energy sources by predicting energy production based on weather patterns, demand, and other factors. This can help ensure that renewable energy sources are used as efficiently as possible.
4. Carbon capture and storage: AI can be used to optimize carbon capture and storage technologies, which can help reduce greenhouse gas emissions by capturing and storing carbon dioxide from power plants and other sources.
5. Sustainable agriculture: AI can be used to optimize agricultural practices, such as crop selection, irrigation, and fertilizer use, to reduce greenhouse gas emissions and increase food production.
6. Climate modeling: AI can be used to improve climate modeling, enabling scientists to better understand the impacts of climate change and develop more effective strategies for mitigating its effects.

In conclusion, AI technology has enormous potential to help address climate change by optimizing energy consumption, managing smart grids, optimizing renewable energy sources, facilitating carbon capture and storage, improving agricultural practices, and enhancing climate modeling. By harnessing the power of AI technology, we can take bold action to address climate change and create a more sustainable future for all."}

5 {"text": "What are some potential benefits of using AI in healthcare? AI has the potential to revolutionize healthcare by helping to improve patient outcomes, reducing costs, and increasing efficiency. Here are some potential benefits of using AI in healthcare:

1. Improved Diagnostics: AI can analyze large amounts of patient data, including medical records, lab results, and imaging scans, to help healthcare professionals diagnose diseases more accurately and quickly. This can lead to earlier diagnoses, more personalized treatments, and better patient outcomes.
2. Personalized Medicine: AI algorithms can analyze patient data to develop personalized treatment plans based on an individual's medical history, genetic makeup, and other factors. This could lead to more effective treatments, fewer side effects, and better patient outcomes.
3. Predictive Analytics: AI can analyze patient data to predict health outcomes and identify patients at risk of developing certain conditions. This could help healthcare professionals intervene early and prevent disease progression, leading to better patient outcomes and reduced healthcare costs.
4. Efficient Resource Allocation: AI can help healthcare organizations allocate resources more efficiently. This includes optimizing patient flow, scheduling appointments, and managing staffing levels, leading to reduced wait times, improved access to care, and better patient outcomes.
5. Drug Discovery: AI can analyze large amounts of data to identify potential new drug candidates and accelerate the drug discovery process. This could lead to the development of new treatments for a wide range of conditions, improving patient outcomes and reducing healthcare costs.
6. Remote Monitoring: AI can be used to remotely monitor patients, allowing healthcare professionals to track patient health and intervene early if necessary. This could lead to improved patient outcomes and reduced healthcare costs by reducing the need for hospitalization.

In conclusion, AI has the potential to bring significant

```
benefits to healthcare by improving patient outcomes, reducing costs, and increasing efficiency. These benefits could lead to better patient experiences, improved access to care, and a more sustainable healthcare system."}
```

预训练的数据集：pretrain_data.csv

</>

Bash

- 1 text
- 2 "我们生产的食品消泡剂，具有可以快速消除泡沫的特点。
- 3 丹东食品消泡剂相关内容：一般而言，纯水和纯表面活性剂不起泡，这是因为它们的表面和内部是均匀的，很难形成弹性薄膜，即使形成亦不稳定，会瞬间消失。
- 4 丹东食品消泡剂选择：
 - 5 1. 相容性：相容性是指两种或者两种以上物质混合时，不产生相斥分离现象的能力，相容性好，消泡剂就能够长期、稳定、均匀地存在于体系中，进而发挥消抑泡的作用；反之，就会出现分层等现象，使消泡剂的消泡工作无法正常进行。
 - 6 2. 消泡能力：消泡能力是消泡剂的最主要性能，鉴别此项性能的标准是在同等条件下，分别加入等量不同的消泡剂，观察消泡剂的消泡速度。"
- 7 "程总在座谈中首先向学校的客人介绍了三一集团和北京三一重机的情况，以及在公司快速发展过程中对人才的渴求，指出通过校企联合，学校可以依靠企业的参与制定人才培养方案，使培养的人才更贴近市场，贴近企业，又可以借助企业的资源充实学校的办学实力。同时校企联合有利于企业的可持续发展。校企联合是企业实现人才战略的途径。企业在与高等职业教育合作过程中可以贯彻自己的培养意向，满足对生产第一线实用型人才的需求。
- 8 武汉交通职业学院盛建龙院长和河北工业职业技术学院李军锁副院长分别介绍了各自学校人才培养情况，并对三一集团的高速发展表示钦佩和赞赏，表示将和公司开展深入、全面的合作，优势互补，使学校和企业实现充分的资源共享，建立全方位长效合作机制。
- 9 本次联合办学签约仪式，是北京桩机高起点校企合作的开始。按照北京桩机人力资源提升计划，明年北京桩机将和所高职高专院校进行联合办学成立“三一班”，均为统招大专高技学历层次，涉及焊接、装配、机加工、售后服务等紧缺工种，“三一班”学员将达到近300人，为北京桩机的下一个五年跨越式发展打下良好的人才基础。"
- 10 "白癜风病人调节心理要偶尔也要屈服。能屈能伸，能进能退，轻松自如；凡事认真，一味固执，肯定烦恼重重。其实，只要大前提不受影响，一些细枝末节上的让步、妥协，是明智的，也是一种大智若愚的姿态。

单轮监督微调的数据集：sft_data_single.csv

</>

Bash

- 1 history,q,a
- 2 [],好的。现在请你将这个文本中的所有逗号都替换成空格。,"好的，请稍等一下，现在我会将文本中的所有逗号替换为空格。处理后文本为："这是一个句子 目的是看看是否可以正确地从这个句子中删除关键词。"。处理结果如何？"
- 3 [],帮我回答一道历史题目。清朝时期的八旗共有多少旗人？,清朝时期八旗旗人总数约为200万人左右，其中正黄旗、正蓝旗、正白旗、正红旗的人数较多，其他旗的人数较少。
- 4 [],请给我介绍一下人工智能的技术和发展。,"人工智能的技术发展已经越来越快，包括深度学习、机器学习、自然语言处理、计算机视觉等。随着算力、数据量和算法的不断提高，人工智能也在不断地进步和演变。未来，人工智能将会被广泛运用于各个领域，包括医疗、金融、智能制造等。与此同时，人工智能也带来了伦理、隐私等方面的问题，需要我们权衡好利弊，做出正确的决策。
- 5 [],很好，现在请你计算一下这篇文章中出现了多少个“获得”字样。,"经过计数，这篇文章中总共出现了两次“获得”这个词。

- 6 ,A公司去年亏损了500万美元,今年净利润增长了50%,今年的净利润是多少? ,今年净利润为750万美元
- 7 ,我也觉得非常棒。最喜欢的是剧中人物的情感变化,尤其是女主角的成长。 ,是的,女主角的发展很让人感动。总的来说,《桥上的人》是一部非常出色的作品。
- 8 ,请从这篇文章中提取出关于分类垃圾的解决方案。 ,文章中提到的第二个方面——分类垃圾,是解决垃圾污染的有效措施之一。垃圾分类是指将生活垃圾按照不同的种类进行分类,以便回收、处理和再利用。通过分类垃圾,我们能够提高废弃物的资源利用率,并减少不必要的浪费,从而保护环境和生态系统。
- 9 ,根据以下文本,对此事件进行分类: 中国队在足球比赛中赢得了冠军。 ,这个事件可以被分类为体育比赛。具体地,中国队在足球比赛中致胜并赢得了冠军。由此可以推断出,这场比赛可能是由各国派出的足球队参加的,中国队在比赛中表现出色并赢得了锦标。 这个事件的重要性对于中国足球运动员和球迷以及全国范围的足球运动发展都具有重要意义。
- 10 ,请告诉我什么是机器学习。 ,机器学习是一种人工智能技术,它使用算法和数学模型,让计算机自动地从数据中学习,并根据这些数据做出预测和决策。机器学习可以被应用在许多不同领域,如金融、医疗保健、广告和交通等。机器学习的目标是为我们提供更好的决策和更好的结果,同时提高效率和减少错误。

多轮监督微调的数据集: sft_data_multi.csv

</>

Bash

- ```
1 history,q,a
2 "[['写一篇关于环境保护的文章。', '好的,请问您需要具体写哪方面的环境保护呢?例如空气质量,水污染等。'], ['写关于减少垃圾污染的文章。', '垃圾污染严重影响环境和我们的健康。要减少垃圾污染,需要有全社会的共同努力。我们可以从以下几个方面入手:\n1.减少塑料袋和一次性容器的使用\n2.分类垃圾,便于回收利用\n3.鼓励使用可生物降解的产品\n4.提高公众的环保意识。我们每个人都可以为保护环境做点力所能及的事情,比如少开车、多步行或骑车上上班。']]",请从这篇文章中提取出关于分类垃圾的解决方案。 ,文章中提到的第二个方面——分类垃圾,是解决垃圾污染的有效措施之一。垃圾分类是指将生活垃圾按照不同的种类进行分类,以便回收、处理和再利用。通过分类垃圾,我们能够提高废弃物的资源利用率,并减少不必要的浪费,从而保护环境和生态系统。
3 "[['请给我生成一首优美的诗歌。', '在苍穹之上,雄鹰在安静中飞行,四周无人,只有欢乐的气息。万物都在静静地生长,却唯独宁静的空气躲藏起来,因为那是一个避难所的名字。'], ['写一篇关于旅行的文章。', '旅行是一种体验,是一种探索未知的方式。你可以在旅途中认识新朋友,了解不同的文化,在神秘的山脉上漫步,到迷人的海滩上玩耍。旅行是一次放松的机会,可以摆脱日常生活中的压力,感受到不同的生活方式。'], ['请帮我想一些头脑风暴的点子。', '好的,让我们一起来头脑风暴吧。我们可以考虑开一家独特的餐厅,或者推出一款创新的手机应用程序。我们还可以尝试打造一个绿色、可持续的产品或服务,或者发现一个新的市场机会。还有,我们可以探索可能的新技术或发明新的产品,以改善人们的生活质量。'], ['请告诉我关于先锋公司的历史和发展情况。', '先锋公司成立于1978年,是一家美国跨国金融服务公司,总部位于宾夕法尼亚州的芝加哥。该公司提供一系列金融产品和服务,包括投资、退休计划、保险和房地产。自成立以来,先锋公司一直致力于为客户提供高效、可靠、低成本的金融服务。它是美国最大的互惠基金管理公司之一,也是世界上最大的投资管理公司之一。']]",请告诉我什么是机器学习。 ,机器学习是一种人工智能技术,它使用算法和数学模型,让计算机自动地从数据中学习,并根据这些数据做出预测和决策。机器学习可以被应用在许多不同领域,如金融、医疗保健、广告和交通等。机器学习的目标是为我们提供更好的决策和更好的结果,同时提高效率和减少错误。
4 "[['请给我写一首关于夏天的诗歌。', '夏天的花儿竞相开放,阳光灿烂,树叶绿茵茵。海风拂过,吹过脸颊,沁人心脾。天空湛蓝,云朵悠悠,一派和谐。这就是夏天,一年四季中最温暖,最绚丽,最有生命力的季节。'], ['请给我列出关于瑜伽的10个好处。', '瑜伽可以缓解身心疲劳,改善睡眠质量,提高身体灵活性和协调性,缓解压力和焦虑,改善心理健康,调节内分泌系统,增强身体免疫力,促进身体健康,减轻病痛,使人更加平静和自我控制。'], ['请告诉我一些当代最流行的音乐类型。', '当代最流行的音乐类型包括流行音乐,电子音乐,摇滚乐和嘻哈音乐。其中,流行音乐和嘻哈音乐受到年轻人的青睐,电子音乐和摇滚乐则吸引了许多热爱音乐的人。但是,音乐的类型和流行趋势会随着时间的推移而不断变化。']]",请解释一下什么是人工
```

智能。人工智能是一种模拟人类智能的技术和方法。它的发展包括机器学习、自然语言处理、计算机视觉等技术的应用。人工智能的目标是让机器可以像人类一样思考、感知、推理和解决问题。例如，人工智能可以帮助我们开发更加智能化的产品，提高生产效率，简化操作流程，优化规划和管理等方面提供帮助。

- 5 "[['根据给定的食材，编写一道美食菜谱。']，['鸡腿肉、蘑菇、洋葱']，['鸡肉炒蘑菇洋葱']\n食材：\n- 鸡腿肉 400克，去骨去皮切成小块\n- 蘑菇 200克，切片\n- 洋葱 1个，切丝\n- 生姜 适量，切丝\n- 酱油 2勺\n- 料酒 1勺\n- 盐 适量\n- 白胡椒粉 适量\n- 植物油 适量\n做法：\n1. 在一个碗里放入切好的鸡腿肉块，加入酱油、料酒和适量的白胡椒粉，搅拌均匀，腌制10分钟。\n2. 热锅凉油，加入姜丝翻炒，放入腌好的鸡肉块翻炒至变色。\n3. 加入洋葱丝和蘑菇，翻炒2-3分钟至蘑菇变软。\n4. 加入适量的盐和少量的水，盖上锅盖焖煮5分钟直至鸡肉熟透。\n5. 最后翻炒几下，使汤汁变浓稠即可。]\n温馨提示：\n1. 在腌制鸡肉时，如果想让肉质更嫩滑，可以加入少量的淀粉和蛋清。]\n2. 可以根据个人口味加入其他调料，如蒜末、辣椒等。]\n3. 洋葱和蘑菇可以用其他蔬菜代替，如青椒、胡萝卜等。']，['太好了，我会试着照着你教的做一次，但是我最近买了一部新相机，想让你推荐一些好的拍照地点。']，['当地的风景区、公园、景点等等都是比较适合拍照的地点。如果你想要一些特别的地方，可以考虑一些建筑群，街道巷子、咖啡馆、书店等具有文艺气息的地方。']]"，嗯，这些地方都很不错，但是我还想找一些可以体验当地文化和风俗的地点，你有什么好的建议吗？，如果你想要更深入地了解当地文化和风俗，可以考虑参加一些游学活动，比如参观当地的博物馆、参加民俗节庆、尝试当地美食等等。此外，也可以去一些小城镇，感受那里的古朴风情。
- 6 "[['写一篇关于人工智能的散文。']，['人工智能是一种前沿的技术，它已经在我们的生活中扮演着越来越重要的角色。对于一些人来说，人工智能是一种奇妙的力量，可以为他们带来更好的生活、更高的效率和更丰富的体验；而对于另外一些人来说，人工智能可能是一种令人担忧的存在，令人担心人工智能最终将夺取人类的地位，毁灭我们的文明。']\n尽管对于人工智能的看法存在重大分歧，但它无疑是当今最具前景和潜力的领域之一。无论是在医疗保健、金融、交通还是安全等领域，人工智能都可以发挥重要的作用，并提供创新的解决方案。]\n不过，人工智能也存在一些潜在的风险和挑战，如人工智能算法的不可解释性、智能系统的失灵和对个人隐私的威胁等。因此，我们需要采取切实有效的措施来应对这些挑战，确保人工智能的可持续和安全发展。]\n总体来说，人工智能是一项极具挑战和前途的领域。随着技术的不断发展和创新，我们相信，未来人工智能将继续发展，并为人类带来更广阔的未来。']]"，请简要概括一下这篇文章的主旨。]，这篇散文主要讨论了人工智能的发展及其可能面临的挑战和风险。尽管人工智能在各个领域中具备着极高的应用价值，但应该认真面对其所带来的挑战，并采取切实有效的措施来确保其可持续和安全发展。
- 7 "[['写一篇关于汉服的文章。']，['好的，请稍等片刻，我会为您生成一篇关于汉服的文章。文章如下：']\n汉服，又称汉衣、华服，是中国传统服饰文化中的一种，源自汉朝，至今已有两千余年的历史。汉服在中国的服饰文化中具有极高的地位和影响力，自古以来被视为自然之美与和谐之美的化身。]\n汉服以它的优美和庄重而闻名于世。古代的汉服分为男女不同的款式和品质，以高质的材料和高超的技艺制成。汉服有着独特的设计和纽扣，以及在服饰上绣着各种图案。同时，汉服也具有不同的场合和用途，从而形成了多种丰富的汉服文化。]\n在现代社会，许多人在尝试重建汉服文化，并将这种传统文化推向更广泛的市场。无论是在电视剧、电影、还是在时装周上，汉服都占据着重要的位置，并备受广大年轻人的欢迎。汉服的崛起也预示着更多传统文化的复兴。]\n如何把汉服推向世界，让更多外国人认识和接受中国的传统文化，已经成为了一个全球性的话题。汉服作为一种传统的服饰文化，有着应该让更多人品味、传承的价值。相信未来，汉服文化一定会继续发扬光大。']]"，请从文章中提取汉服的历史。]，汉服是中国传统服饰文化中的一种，源自汉朝，至今已有两千余年的历史。

dp

## 1. 预训练 (Pretain)

### 1.1 文本如何变成embedding的索引数组和embedding张量

在下一步就要进行分词时数据集的样子：

</>

Bash

- 1 <s>"我们生产的食品消泡剂，具有可以快速消除泡沫的特点。</s>
- 2 <s>丹东食品消泡剂相关内容：一般而言，纯水和纯表面活性剂不起泡，这是因为它们的表面和内部是均匀的，很难形成弹性薄膜，即使形成亦不稳定，会瞬间消失。</s>
- 3 <s>丹东食品消泡剂选择：</s>

## 在转变为embedding的索引数组后：

x在这里是一个整数，是词表的索引

</>

Bash

- 1 x x x x x x x x
- 2 x x x x
- 3 x x x x x x x
- 4 x x x x x x x

## 在转变为embedding的索引数组且padding后：

索引数组都小于需要的最大的长度时，补零。0在词表里是<unk>：

</>

Bash

- 1 x x x x x x x x 0 0 0 0
- 2 x x x x 0 0 0 0 0 0 0 0
- 3 x x x x x x x 0 0 0 0 0
- 4 x x x x x x x 0 0 0 0 0

索引数组的shape应该为：[batch\_size,max\_length]

索引数组通过embedding表查询到embedding后获得的张量大小为：[batch\_size,max\_length, embding\_size]

## 索引数组还伴随一个loss掩码数组：

</>

Bash

- 1 1 1 1 1 1 1 1 1 0 0 0 0
- 2 1 1 1 1 0 0 0 0 0 0 0 0
- 3 1 1 1 1 1 1 1 0 0 0 0 0
- 4 1 1 1 1 1 1 1 0 0 0 0 0

loss掩码数组用来mask掉padding的0的loss值计算，在反向传播时不传播它们的loss或者让loss为零不起作用。



## 构造输入与监督输出:

```
</> Bash

1 原始分词后的文本: <s> token1 token2 token3 token4 token5 </s> 0 0 0 0 0
2 模型输入为 (去掉第一个token): <s> token1 token2 token3 token4 token5 </s>
 0 0 0 0
3 模型监督输出为 (去掉最后一个token): token1 token2 token3 token4 token5 </s> 0
 0 0 0 0
4 输入与监督输出错位是为了有:
5 <s> -----> 预测 token1
6 <s> token1 -----> 预测 token2
7 <s> token1 token2 -----> 预测 token3
8 <s> token1 token2 token3 -> 预测 token4
```

## 1.2 执行单机四卡训练

执行命令:

```
</> Bash

1 nohup torchrun --nproc_per_node 4 1-pretrain.py > pretrain.log 2>&1 &
```

执行日志:

```
</> Bash

1 [2025-01-21 08:06:43,125] torch.distributed.run: [WARNING]
2 [2025-01-21 08:06:43,125] torch.distributed.run: [WARNING]

3 [2025-01-21 08:06:43,125] torch.distributed.run: [WARNING] Setting OMP_NUM_THREADS
 environment variable for each process to be 1 in default, to avoid your system
 being overloaded, please further tune the variable for optimal performance in your
 application as needed.
4 [2025-01-21 08:06:43,125] torch.distributed.run: [WARNING]

5 LLM总参数量: 26.878 百万
6 Epoch:[0/20](0/20957) loss:8.879 lr:0.0002000 epoch_Time:1372.0min:
7 Epoch:[0/20](100/20957) loss:7.420 lr:0.0002000 epoch_Time:84.0min:
8 Epoch:[0/20](200/20957) loss:6.885 lr:0.0002000 epoch_Time:76.0min:
9 Epoch:[0/20](300/20957) loss:6.428 lr:0.0002000 epoch_Time:73.0min:
10 ...
11 Epoch:[19/20](20200/20957) loss:2.264 lr:0.0000200 epoch_Time:2.0min:
12 Epoch:[19/20](20300/20957) loss:2.280 lr:0.0000200 epoch_Time:2.0min:
13 Epoch:[19/20](20400/20957) loss:2.249 lr:0.0000200 epoch_Time:2.0min:
```

```
14 Epoch: [19/20](20500/20957) loss:2.369 lr:0.0000200 epoch_Time:1.0min:
15 Epoch: [19/20](20600/20957) loss:2.195 lr:0.0000200 epoch_Time:1.0min:
16 Epoch: [19/20](20700/20957) loss:2.399 lr:0.0000200 epoch_Time:1.0min:
17 Epoch: [19/20](20800/20957) loss:2.339 lr:0.0000200 epoch_Time:0.0min:
18 Epoch: [19/20](20900/20957) loss:2.391 lr:0.0000200 epoch_Time:0.0min:
```

文件系统输出:

```
</> Bash
1 ./out/
2 └─ pretrain_512.pth
```

## 1.3 评估模型

执行命令:

```
</> Bash
1 python 0-eval_pretrain.py
```

执行日志:

```
</> Bash
1 模型参数: 26.878464 百万 = 0.026878464 B (Billion)
2 问题: 椭圆和圆的区别
3 回答:
4 椭圆和圆的区别
5 椭圆的区别主要是椭圆的形状和大小, 椭圆的形状和大小, 椭圆的形状和大小不同。椭圆椭圆的圆, 椭圆的形
状, 椭圆的形状, 椭圆的形状。椭圆的大小, 椭圆的形状和大小, 椭圆的形状, 椭圆的形状和大小。椭圆的形状
和大小, 椭圆的形状和大小, 椭圆的形状和形状, 椭圆的形状, 椭圆的大小和大小, 椭圆的
形状和形状, 椭圆的形状, 椭圆的形状和大小, 椭圆的形状和形状。
6 椭圆的形状和大小
7 椭圆的形状, 椭圆的形状和形状, 椭圆的形状和形状是椭圆的形状, 椭圆的形状和形状为椭圆的形状。椭圆
的形状和形状, 椭圆的形状和形状是椭圆的形状和形状。椭圆的形状和形状。椭圆的形状和形状, 椭圆的形状和形状
是椭圆的形状。椭圆的形状和形状为椭圆的形状。椭圆的形状和形状。椭圆的形状和形状是椭圆的形状。椭圆
的形状和形状为椭圆的形状和形状为椭圆的形状和形状为椭圆的形状和形状为椭圆的形状和形状为椭圆的形状。
8 椭圆的形状和形状是椭圆的形状和形状, 椭圆的形状和形状是椭圆的形状和形状。椭圆是椭圆的形状和形状, 椭
圆是椭圆, 椭圆是椭圆的形状。椭圆的形状和形状都是椭圆的形状和形状。椭圆的形状和形状是椭圆的形状和形
状。
9
10 3.762782573699951 s
11 问题: 中国关于马克思主义基本原理
12 回答: 概论的论述
```



35 世界上人口最少的国家是：墨西哥、美国、墨西哥、中国、日本、澳大利亚、新西兰、澳大利亚和新西兰。

36 世界上人口最少的国家是：墨西哥、美国、墨西哥、中国、日本、澳大利亚和新西兰。

37 世界上人口最少的国家是：墨西哥、美国、墨西哥、中国、日本、澳大利亚和新西兰。

38 世界上人口最少的国家是：墨西哥、美国、墨西哥、中国、日本、澳大利亚、新西兰。

39 世界上人口最少的国家是：墨西哥、美国、墨西哥、中国、日本、澳大利亚和新西兰。

40

41 2.0161516666412354 s

42 问题： DNA的全称是

43 回答： DNA，是通过丈夫与妻子之间的亲缘关系而形成的，是亲缘关系。DNA的作用包括：

44 1、 将一个具有DNA序列的遗传基因，将一个具有一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将其移植到一个具有DNA序列的遗传基因上。

45 2、 通过将具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个具有DNA序列的遗传基因，将一个属于第三种遗传性遗传基因的遗传基因，将一个属于第四种遗传基因，将一个属于第四种遗传基因，将一个属于第四种遗传基因。

46

47 2.8412325382232666 s

48 问题： 数学中 $\pi$ 的值大约是

49 回答： 1.5，但是在数学中 $\pi$ 的值是1.7，所以数学中 $\pi$ 的值在数学中占的比例很小。

50  $\pi$ 中 $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值是1.5。在数学中 $\pi$ 的值是1.5。 $\pi$ 中 $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值是1.5。

51  $\pi$ 中的 $\pi$ 中 $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值是1.5。

52  $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值在数学中占的比例很小。

53  $\pi$ 中 $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值在数学中占的比例很小。

54  $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值在数学中占的比例很小。

55  $\pi$ 中 $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值在数学中占的比例很小，所以数学中 $\pi$ 的值在数学中占的比例很小。

56 以上就是关于数学中 $\pi$ 的值在数学中占的比例，希望对大家有所帮助。

57

58 2.6969528198242188 s

59 问题： 世界上最高的山峰是

60 回答： 西湖，它在西湖上有三座山峰，它位于山顶。山峰高约15米，宽1.5米，高1.85米，峰顶长约2.95米，峰顶有4座山峰，峰顶有1座峰，峰顶有5个岩石，山顶有2个岩石，峰顶有4个岩石，峰顶有一座山峰。山峰高约15米，峰顶有1座山峰，峰顶有1座山峰，峰顶有1座山峰，峰顶有3座山峰。山峰高约2.4米，峰顶有1座山峰，峰顶有1座山峰。山峰高约15米，峰顶有5个岩石，峰顶有1座山峰，峰顶有2座山峰。山峰高约15米，峰顶有2座山峰，峰顶有1座山峰。峰顶有1座山峰。山峰高约15米，峰顶有1座山峰。峰顶有4个岩石，峰顶有4个岩石，峰顶有1座山峰。山峰高约15米，峰顶有1座山峰。山高约2米，峰顶有3座山峰。峰顶有1座山峰，峰顶有1座山峰。峰顶有1座山峰，峰顶有1座山峰。峰顶有4座山峰，峰顶有4座山峰。峰顶有3座山峰。峰顶有1座山峰。峰顶有1座山峰。峰顶有1座山峰。峰顶有1座山峰。峰顶有2座山峰。峰顶有1座山峰。峰顶有1座山峰。峰顶有2个岩

石，峰顶有2座山峰。峰顶有1座山峰。峰顶有1座山峰。峰顶有1座山峰。峰顶有1座山峰。峰顶有2座山峰，峰顶有2座山峰。峰上有2座山峰。峰顶有1座山峰。峰顶有1座山峰，峰顶有1座山

61

62 3.764134645462036 s

63 问题： 太阳系中最大的行星是

64 回答：地球。

65 地球是地球的星体，太阳系最大的行星是地球。

66 地球是地球。

67 地球是地球。

68 太阳系中的地球，是地球。

69 宇宙中，地球是地球。

70 地球是地球。

71 世界上，地球是地球。

72 地球是地球。

73 地球是地球。

74 地球是地球。

75 地球是地球。

76 地球是地球。

77 地球是地球。

78 地球是地球。

79 地球是地球。

80 地球是地球。

81 地球是地球。

82 地球是地球。

83 地球是地球。

84 地球是地球。

85 地球是地球。

86 地球是地球。

87 地球是地球。

88 地球是地球。

89 地球是地球。

90 地球是地球。

91 地球是地球。

92 地球是地球。

93 地球是地球。

94 地球是地球。

95 地球是地球。

96 地球是地球。

97 地球就是地球。

98 地球是地球。

99 人类都是地球。

100 地球是地球。

101 地球是地球。

102 地球是地球。

103 地球是地球。

104 地球是地球。

105 地球是地球。

106 地球是地球。

107 地球是地球。

108 地球是地球。

109 地球是地球。

110

111 1.6594161987304688 s

112 问题： 二氧化碳的化学分子式是

113 回答：CO<sub>2</sub>，分子量为590，是一种重要的能源，在石油工业中是不可缺少的能源之一。二氧化碳的化学分子式是CO<sub>2</sub>，分子量为590，是一种重要的能源之一。在石油工业中，二氧化碳是一种重要的化工原料，二氧化碳具有极强的抑制和分解作用，是重要的工业原料之一。

114

115 0.5974256992340088 s

116 问题： 地球上最大的动物是

117 回答：水母。地球上最大的动物是水母。它是地球上唯一的动物。

118 地球上最大的动物是水母。它是地球上唯一的动物。它是地球上唯一的动物。

119 地球上最大的动物是水母。它是地球上唯一的动物。它是地球上唯一的动物。

120 地球上最大的动物是水母。它是地球上唯一的动物。

121

122 0.6156930923461914 s

123 问题： 地球自转一圈大约需要

124 回答：十六分钟，而地球自转一圈大约需要十五分钟，而地球自转一圈大约需要十七分钟。

125 地球自转一圈大约需要十五分钟，而地球自转一圈大约需要十八分钟，而地球自转一圈大约需要十七分钟。

126 地球自转一圈大约需要十八分钟，而地球自转一圈大约需要十九分钟。

127 地球自转一圈大约需要十八分钟。

128 地球自转一圈大约需要十七分钟，而地球自转一圈大约需要十九分钟。

129

130 0.9954221248626709 s

131 问题： 杭州市的美食有

132 回答：限公司秉承“质量为本，服务社会”的原则，立足于高新技术，科学管理，拥有现代化的生产、检测及试验设备，已建立起完善的产品结构体系，产品品种，结构体系完善，性能质量稳定。

133 杭州市的美食有限公司是一家具有完整生态链的企业，它为客户提供综合的、专业现代化装修解决方案。为消费者提供较优质的产品、较贴切的服务、较具竞争力的营销模式。

134

135 0.8061814308166504 s

136 问题： 江苏省的最好的大学

137 回答：

138 大学

139 江苏大学，是经江苏省人民政府批准设立的全日制普通本科高校，也是江苏省唯一一所独立设置的公办普通本科高校。江苏省的最大的高校是南京大学，在江苏地区的高校中排名第四，江苏省高校中排名第四，江苏省高校中排名第六，江苏省高校中排名第七。

140 江苏省的最好的大学



```
</s>"
```

## minimind的chat\_template:

位于: minimind/model/minimind\_tokenizer/tokenizer\_config.json:43

这是一个类似Jinja的模版:

```
</> Bash

1 {% if messages[0]['role'] == 'system' %}
2 {% set system_message = messages[0]['content'] %}
3 {% endif %}
4 {% if system_message is defined %}
5 {{ system_message }}
6 {% endif %}
7 {% for message in messages %}
8 {% set content = message['content'] %}
9 {% if message['role'] == 'user' %}
10 {{ '<s>user\\n' + content + '</s>\\n<s>assistant\\n' }}
11 {% elif message['role'] == 'assistant' %}
12 {{ content + '</s>' + '\\n' }}
13 {% endif %}
14 {% endfor %}
```

一个简单的实例(空格也是格式的一部分):

```
</> Bash

1 你是小明, 家住山的那边
2 <s>user\\n
3 你叫什么名字? </s>\\n
4 <s>assistant\\n
5 我叫小明</s>\\n
6 <s>user\\n
7 你家住哪里? </s>\\n
8 <s>assistant\\n
9 我家住在那边的山上</s>\\n
```

变成一行输入:

```
</> Bash

1 你是小明, 家住山的那边<s>user\\n你叫什么名字? </s>\\n<s>assistant\\n我叫小明
 </s>\\n<s>user\\n你家住哪里? </s>\\n<s>assistant\\n我家住在那边的山上</s>\\n
```

变成一行后, 就是预训练中类似的一条普通训练样本



单轮对话是指：没有system\_message,只有一个user\_message和一个assistant\_message

## 2.2 执行单机四卡单轮对话全参数监督微调

执行命令：

```
</> Bash

1 torchrun --nproc_per_node 4 3-full_sft.py
```

执行日志：

```
</> Bash

1 nohup: ignoring input
2 [2025-01-22 11:03:08,346] torch.distributed.run: [WARNING]
3 [2025-01-22 11:03:08,346] torch.distributed.run: [WARNING]

4 [2025-01-22 11:03:08,346] torch.distributed.run: [WARNING] Setting OMP_NUM_THREADS
 environment variable for each process to be 1 in default, to avoid your system
 being overloaded, please further tune the variable for optimal performance in your
 application as needed.
5 [2025-01-22 11:03:08,346] torch.distributed.run: [WARNING]

6 LLM总参数量： 26.878 百万
7 Epoch:[0/19](0/30851) loss:2.919 lr:0.0001000 epoch_Time:813.0min:
8 Epoch:[0/19](100/30851) loss:2.256 lr:0.0001000 epoch_Time:63.0min:
9 Epoch:[0/19](200/30851) loss:2.320 lr:0.0001000 epoch_Time:60.0min:
10 Epoch:[0/19](300/30851) loss:2.114 lr:0.0001000 epoch_Time:58.0min:
11 ...
12 Epoch:[15/19](26400/30851) loss:1.541 lr:0.0000159 epoch_Time:8.0min:
13 Epoch:[15/19](26500/30851) loss:1.680 lr:0.0000159 epoch_Time:8.0min:
14 Epoch:[15/19](26600/30851) loss:1.248 lr:0.0000159 epoch_Time:8.0min:
15 Epoch:[15/19](26700/30851) loss:1.580 lr:0.0000159 epoch_Time:7.0min:
16 Epoch:[15/19](26800/30851) loss:1.407 lr:0.0000159 epoch_Time:7.0min:
17 Epoch:[15/19](26900/30851) loss:1.783 lr:0.0000159 epoch_Time:7.0min:
```

## 2.3 评估模型的单轮对话

执行命令：

```
</> Bash

1 python 2-eval.py
```

1 模型参数：26.878464 百万 = 0.026878464 B (Billion)

2 [Q]：你叫什么名字

3 [A]：作为一个虚拟的AI，我没有身体，所以不需要叫什么名字。但是，我可以回答你的问题，提供信息和建  
议，帮助你解决问题。

4

5 [Q]：你是谁

6 [A]：我是一个由人工智能技术构建的语言模型，没有实际的人类情感和思维能力。

7

8 [Q]：中国有哪些比较好的大学？

9 [A]：中国有很多优秀的大学，比如上海上海交通大学、深圳的国际学校、广州的交通大学、深圳的科技大学  
等等。

10

11 [Q]：全世界最好的大学是什么？

12 [A]：这个问题很难回答，因为每个人对大学的定义都不尽相同。但是，根据一些大学的声誉和排名，一些大学  
被认为是最好的大学。例如哈佛大学、斯坦福大学等。

13

14 [Q]：你知道光速是多少吗？

15 [A]：光速是每秒约30万公里或 $2 \times 35,995,000$ 米的速度。它是物理学中最重要和最广泛的量之一，被广泛应用于  
通讯、地球物理学、天体物理学、计算机科学、天文学等领域。

16

17 [Q]：你知道长江吗？

18 [A]：当然，长江是中国的母亲河，长江流经中国多个省份和地区，是中国历史文化的重要发源地之一。

19

20 [Q]：人类的血液主要由哪些成分组成？

21 [A]：人类的血液主要由葡萄糖和脂肪酸构成。葡萄糖是身体的主要能量来源，而脂肪酸则是身体内许多重要的  
营养素的来源。

22

23 [Q]：第一颗人造卫星是哪个国家发射的？

24 [A]：第一颗人造卫星是苏联的斯普特尼克一号，它于1957年成功地将苏联送入了轨道，并于1960年成功着陆  
在月球上。

25

26 [Q]：你知道杭州有什么美食吗？

27 [A]：杭州的美食非常多样化，其中西湖醋鱼、龙井虾仁、东坡肉、东坡肉等都是非常出名的美食。此外，杭州  
还有很多特色小吃，比如西湖醋鱼、龙井虾仁、叫化鸡等。

28

29 [Q]：你知道泰山在哪里吗？

30 [A]：当然，泰山位于中国山东省中部，是中国五岳之一，也是历史悠久的名山之一。

31

32 [Q]：地球上最大的动物是什么？

33 [A]：地球上最大的动物是蓝鲸，体重可达到200吨以上，长度可达到30米以上。

34

35 [Q]: 地球自转一圈大约需要多少时间?

36 [A]: 地球自转一圈大约需要365.24天左右, 这被称为地球的自转周期。地球自转周期是23小时56分钟, 也就是大约43小时56分钟, 这也被称为地球的自转周期。

37

38 [Q]: 人类最早使用的金属是什么?

39 [A]: 人类最早使用的金属是石质, 据科学家研究, 石英可以被用于制造更轻便的机器和设备。

40

41 [Q]: 水的化学分子式是什么?

42 [A]: 水的化学式为H<sub>2</sub>O, 表示每两分子由两个氢原子和一个氧原子组成。

43

44 [Q]: 大气层中含量最多的气体是什么?

45 [A]: 大气层中含量最多的气体是氮气, 占比约为78%。其他气体包括氩气、二氧化碳、氦气、氟气、氙气和少量的氡气, 它们都是大气层中占比最多的气体。

46

47 [Q]: 世界上最高的山峰是什么?

48 [A]: 世界上最高的山峰是珠穆朗玛峰, 位于喜马拉雅山脉, 海拔8848米。

49

50 [Q]: 你知道世界上最深的海沟是什么吗?

51 [A]: 世界上最深的海沟是马里亚纳海沟, 位于西太平洋, 深度达到11,034米。该海沟是由太阳板和地壳形成的, 其地壳的来势非常强大, 深度超过10公里。

52

53 [Q]: 最早发明印刷术的是哪个国家?

54 [A]: 最早发明印刷术的国家是中国。

55

56 [Q]: 万有引力是谁提出的?

57 [A]: 万有引力是由古典物理学家伽利略提出的。他认为, 所有万有引力都是由于质量产生的, 物体的质量会导致它们之间的相对运动。这种现象是由于物体之间的相互作用所引起的。

58

59 [Q]: 光合作用的主要原理是什么?

60 [A]: 光合作用是植物通过叶绿素等色素吸收太阳光能, 将二氧化碳和水转化成有机物质(如葡萄糖)的过程。在光合作用中, 光能被光合色素吸收, 并转化为有机化合物。这些有机物质是叶绿体中的色素, 在光合作用过程中, 光合产物中的叶绿素吸收了能量, 利用光能转化为ATP和NADPH, 同时释放出能量。这个能量被用于将二氧化碳和水转化成葡萄糖和氧气。光合作用是维持地球大气氧气含量和生命体系的重要过程之一。

61

62 [Q]: 你知道大熊猫的主要食物是什么吗?

63 [A]: 大熊猫主要以竹子为食, 也会吃些水果、坚果、小型哺乳动物、鸟类和爬行动物等。

64

65 [Q]: 海水为什么是咸的?

66 [A]: 海水咸的原因是因为水中的盐分较多, 而咸水在地球上普遍存在的, 同时海水中含有大量的盐分。海水中的盐分主要来自于岩石中的矿物质和微生物, 例如淡水鱼、海藻、海藻、海藻等。这些物质在海水中溶化或分解, 产生了大量的盐分。在咸水的过程中, 这些盐分被带到水中, 形成了咸味的咸味。

67

68 [Q]: 我们平时喝的牛奶主要含有什么营养成分?

69 [A]: 喝牛奶主要含有很多营养成分, 包括蛋白质、脂肪、碳水化合物、维生素、矿物质等。其中, 蛋白质是人体细胞的重要组成部分, 而碳水化合物则是能量的产物, 维生素和矿物质则是身体所需的能量来源。因此, 不同人的喝牛奶的营养成分是不同的。一般来说, 人体需要摄入适量的维生素、矿物质、蛋白质等营养成分, 而成熟的牛奶则需要维生素和矿物质, 因此建议将牛奶和水混合后再食用。

70

71 [Q]: 一星期有多少天?

72 [A]: 一周有七天, 也就是说一周有七天。

从执行日志看, sft训练是成功的。模型学会了一问一答, 而且答案是合理的

文件系统输出:

```
</> Bash

1 ./out/
2 └─ full_sft_512.pth
3 └─ pretrain_512.pth
```

## 3. 多轮对话的全参数监督微调

### 3.1 数据集如何变成embedding的索引数组和embedding数组

参见2.1

多轮对话是指: 可能有system\_message, 有多个user\_message和assistant\_message对儿

### 3.3 执行单机四卡多轮对话全参数监督微调

```
</> Bash

1 TODO
```

## 4. DPO训练

理解参考: <https://zhuanlan.zhihu.com/p/715114620>

### 4.1 DPO损失函数

## DPO 的损失函数

DPO 的损失函数基于 **Bradley-Terry** 模型，这是一种用于成对比较的概率模型。给定一对生成结果  $(y_1, y_2)$ ，其中  $y_1$  是人类偏好的输出， $y_2$  是非偏好的输出，DPO 的损失函数定义如下：

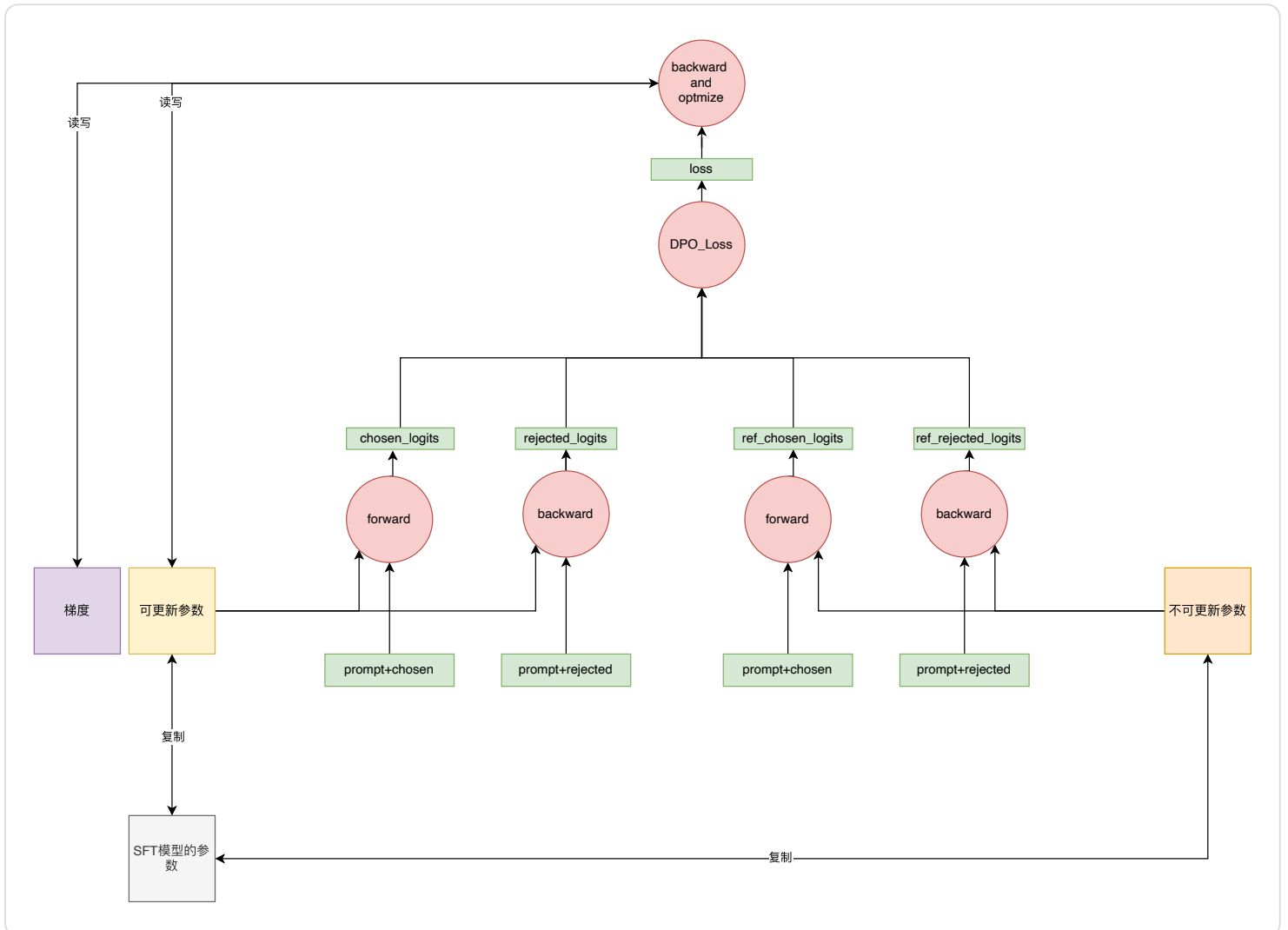
$$\mathcal{L}_{\text{DPO}} = -\mathbb{E}_{(x, y_1, y_2) \sim \mathcal{D}} \left[ \log \sigma \left( \beta \left( \log \frac{\pi_{\theta}(y_1|x)}{\pi_{\text{ref}}(y_1|x)} - \log \frac{\pi_{\theta}(y_2|x)}{\pi_{\text{ref}}(y_2|x)} \right) \right) \right]$$

### 公式解析：

- $\pi_{\theta}(y|x)$ :
  - 待优化的策略模型，参数为  $\theta$ ，表示在输入  $x$  下生成输出  $y$  的概率。
- $\pi_{\text{ref}}(y|x)$ :
  - 参考模型（通常是一个经过 SFT 的模型），用于防止优化后的模型偏离初始模型太远。
- $\log \frac{\pi_{\theta}(y|x)}{\pi_{\text{ref}}(y|x)}$ :
  - 策略模型和参考模型的对数概率比，用于衡量生成结果的相对优劣。
- $\sigma(\cdot)$ :
  - Sigmoid 函数，将差值映射到  $[0, 1]$  之间，表示偏好概率。
- $\beta$ :
  - 温度参数，控制模型对偏好数据的敏感程度。
- $\mathcal{D}$ :
  - 人类偏好数据集，包含输入  $x$  和成对的偏好输出  $(y_1, y_2)$ 。

知乎 @Evan Lee

## 4.2 一步DPO训练



# 4.3 执行DPO训练

TODO