Sentiment情感分析实验 5月6日完成

基于BERT的自动问答实验 5月15日完成

nlpcc2016kbqa

GCN科学出版物分类实验 5月22日完成



昱峰,这是Sentiment情感分析实验在启智平台上已训练跑通过的代码，你看一下

Sentiment情感分析实验在启智平台上已训练跑通过；

基于BERT的自动问答实验需要将数据集上传到启智平台，但通过命令的上传会出现乱码，因此需要通过克隆的方式上传，这就需要先上传到git上，因此需要开通git权限，已经在找王春爱经理申请开通

骆阳，你好，有个bug想请教你，是昱峰推荐的，有个RuntimeError问题，不知道如何解决，网上找了几种方法都不行

Lwx1251421

unzip src.zip 解压文件

ll 查看目录文件

我们需要从官方网站下载NLPCC2016数据集，该数据集包含训练和测试数据。链接为：[http://tcci.ccf.org.cn/conference/2016/pages/page04\_kbqa.html](http://tcci.ccf.org.cn/conference/2016/pages/page04_kbqa.html" \t "_blank)

从NLPCC 2016官网下载相应的KBQA数据集，包括训练集、验证集和测试集。下载链接：<http://tcci.ccf.org.cn/conference/2016/pages/page05_PerformanceEvaluation.html>



昱峰，基于BERT的自动问答实验nlpcc2016kbqa在网上找不到数据集，是否可以使用squad数据集？

然后squad数据集上传到启智总是失败，上传到github上由于有文件大于100M，也上传失败。这个该咋办？

昱峰，我这两天在安装环境，其中win+cpu的mindsore已安装好。

由于linux以前没安装过，所以目前按网上教程安好了VM虚拟机和Ubuntu18.04，

但出现两个问题，一是无法切换到root环境，二是linux连不上网。

目前正在网上找解决方案。

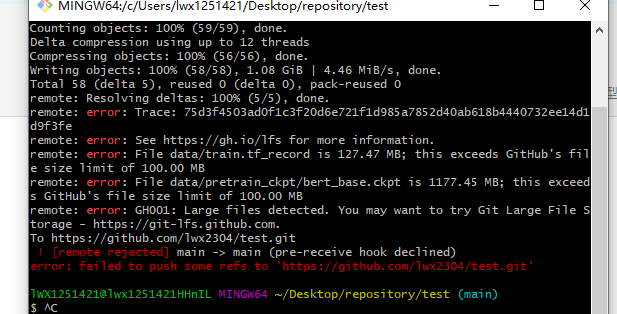
元元总，我这两天在安装环境，其中win+cpu的mindsore已安装好。

由于linux以前没安装过，所以目前按网上教程安好了VM虚拟机和Ubuntu18.04，

但出现两个问题，一是无法切换到root环境，二是linux连不上网。

目前正在网上找解决方案。

元元总，基于BERT的自动问答实验，由于数据集传送到启智平台上总是失败，故在本地win+cpu+mindspore2.0rc1上调试代码，结果出现问题，问了昱峰和骆阳说是不支持。最后通过关闭xGate后，成功将数据集传送到启智平台上。明天开始启智上的调试。



退出git: ctrl+c

git clone https://github.com/lwx2304/test.git/

https://github.com/lwx2304/test.git

[liuweixing5@h-partners.com](mailto:liuweixing5@h-partners.com)

在启智平台上运行Sentiment情感分析实验

1.win+cpu上安装mindspore；安装VM虚拟机和Ubuntu18.04

2、在启智平台上运行Sentiment情感分析实验

1、智能基座深度学习课程图文案例刷新一个。  
2、兰州大学5个实验案例及对应指导书验收，并反馈问题。  
3、更新兰州大学、南京大学周报。  
4、验收南京大学实验案例及实验指导书2个；并解决学校老师在整改过程中遇到的问题。  
5、整理第三期众智案例合同中的验收标准及验收方法。  
进度：  
1、完成100%  
2、完成90%  
3、完成100%  
4、完成95%  
5、完成100%

1.win+cpu上安装mindspore；安装VM虚拟机和Ubuntu18.04

2、在启智平台上运行Sentiment情感分析实验并得到运行结果

进度：  
1、完成100%  
2、完成100%

包括数据下载模块、加载IMDB数据集、加载预训练词向量、数据集预处理、模型构建、损失函数与优化器、训练逻辑、评估指标和逻辑、模型训练与保存、模型加载与测试、自定义输入测试

1.基于BERT的自动问答实验，导入依赖库、运行环境配置和定义超参数

2. 数据预览、定义数据预处理函数和定义训练函数

进度：  
1、完成100%  
2、完成100%

1. 基于BERT的自动问答实验，执行训练、定义评估函数、评估模型效果和单例评估
2. 书写基于BERT的自动问答实验markdown文档

3.西南交大线稿提取项目训练、测试

1. WARNING
2. 在启智NPU环境上运行成都中医院大学数字中医药分析项目
3. 书写成都中医院大学数字中医药分析项目验收报告
4. 西南交大线稿提取项目在CPU上训练和测试

1.基于BERT的自动问答实验，执行训练、定义评估函数、评估模型效果和单例评估

2.书写基于BERT的自动问答实验markdown文档，但有2个WARNING需要消除

3.在启智NPU环境上运行成都中医院大学数字中医药分析项目

4.书写成都中医院大学数字中医药分析项目验收报告

5.西南交大线稿提取项目在CPU上训练和测试

进度：  
1、完成100%  
2、完成100%  
3、完成100%  
4、完成100%

5、完成100%

周六加班产出：修改成都中医院大学数字中医药分析项目验收报告

1. 在启智NPU环境上运行成都中医院大学数字中医药分析项目
2. 书写成都中医院大学数字中医药分析项目验收报告

四川美术学院基于 CycleGAN 算法风格迁移功能进行字体风格生成运算

1.基于BERT的自动问答实验，执行训练、定义评估函数、评估模型效果和单例评估

2.书写基于BERT的自动问答实验markdown文档

3.基于BERT的自动问答实验，消除5个WARNING

1.西南交大线稿提取项目在CPU上训练

2.西南交大线稿提取项目进行测试和交互测试

3.在启智NPU环境上运行成都中医院大学数字中医药分析项目

4.书写成都中医院大学数字中医药分析项目验收报告

进度：  
1、完成100%  
2、完成100%  
3、完成100%  
4、完成100%

西南交大线稿提取项目[Yiwen2022](https://openi.pcl.ac.cn/Yiwen2022) / [CycleGan\_DexiNed](https://openi.pcl.ac.cn/Yiwen2022/CycleGan_DexiNed)

启智NPU上训练、测试和交互测试

推理由于cmake版本升级不了所以不能推理

1. 书写基于BERT的自动问答实验markdown文档

2.基于BERT的自动问答实验，消除5个WARNING

3.西南交大线稿提取项目 [CycleGan\_DexiNed](https://openi.pcl.ac.cn/Yiwen2022/CycleGan_DexiNed)启智NPU上训练、测试和交互测试

4.召集会议评审西南交大线稿提取项目

进度：  
1、完成100%  
2、完成100%  
3、完成100%  
4、完成100%

1.西南交大线稿提取项目在CPU上训练  
2.西南交大线稿提取项目进行测试和交互测试  
3.在启智NPU环境上运行成都中医院大学数字中医药分析项目  
4.书写成都中医院大学数字中医药分析项目验收报告  
进度：  
1、完成100%  
2、完成100%  
3、完成100%  
4、完成100%  
元元总，这周的工作内容

大家以后半周报、月报、工作进度都统一一下格式哈，不然整理起来有点麻烦  
1、跟踪解决论坛问题；  
2、审核CSDN第四期博文；  
3、审核第四期报错活动帖子；  
4、整理奖品发票；  
5、迁移问题总结帖子；

1、启智NPU上cmake版本升级问题解决；

2、西南交大线稿提取项目启智NPU上310推理跑通；

3、西南交大线稿提取项目验收报告书写；

1、西南交大线稿提取项目启智NPU上310推理跑通；

2、西南交大线稿提取项目验收报告书写；

1、和华为员工沟通自动问答数据集上传并返回URL；

2、基于BERT的自动问答提交PR出错，解决报错；

1、和华为员工沟通自动问答数据集上传并返回URL；

2、学习软件工程与技术和数据结构与算法，为可信考试做准备；

基于BERT的自动问答实验：

保证保存和加载时的网络结构一致



如何去掉警告！

通过训练保存模型CKP\_1-3\_2745.ckpt

把"load\_pretrain\_checkpoint\_path":"./data/pretrain\_ckpt/bert\_base.ckpt"替换为

"load\_pretrain\_checkpoint\_path":"./data/pretrain\_ckpt/CKP\_1-3\_2745.ckpt"

这样便可以消除以上警告，使得保存和加载时的网络结构一致。

加载bert\_base.ckpt由于5个参数模型加载不上去，会出现警告，因此需要训练好一个ckpt文件CKP\_1-3\_2745.ckpt用来替换bert\_base.ckpt，这样参数才能都加载上去，没有警告，但CKP\_1-3\_2745.ckpt有1.1G，比较大，在启智NPU调试任务中下载不了。因此在这里说明下。

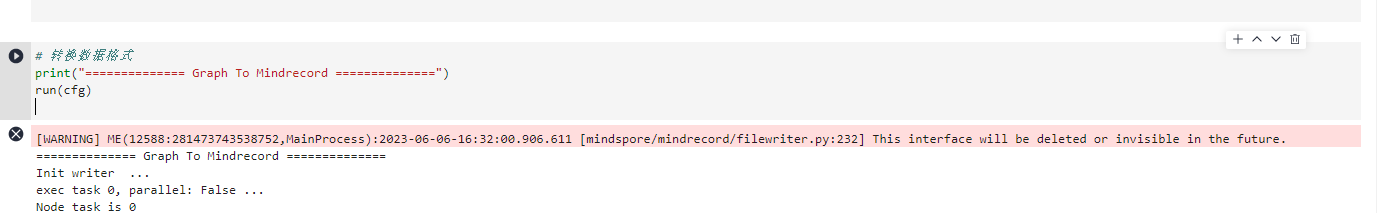
替换如下图所示：



变为：



GCN科学出版物分类：



如何消除：



主要有告警的代码在这里面



直接把这句注释掉，就没问题

1. 书写基于BERT的自动问答实验markdown文档  
2.基于BERT的自动问答实验，消除5个WARNING   
3.西南交大线稿提取项目 CycleGan\_DexiNed启智NPU上训练、测试和交互测试  
4.召集会议评审西南交大线稿提取项目  
进度：  
1、完成100%  
2、完成100%  
3、完成100%  
4、完成100%  
  
刘卫星(wx1251421) 2023-06-09 15:59  
1、西南交大线稿提取项目启智NPU上310推理跑通；   
2、西南交大线稿提取项目验收报告书写；  
3、和华为员工沟通自动问答实验数据集上传并返回URL链接；   
4、学习软件工程与技术和数据结构与算法，为可信考试做准备；  
进度：  
1、完成100%；  
2、完成100%；  
3、完成100%；  
4、完成80%；

AI+x验收，成都中医院大学数字中医药分析项目和西南交大线稿提取项目已验收完毕，验收报告已写；四川美术学院基于 CycleGAN 算法风格迁移正在验证中。

1. 利用download下载数据集在启智NPU运行自动问答实验；
2. 考试可信考试的软件工程与技术和信息安全考试；

1. 四川美术学院基于CycleGAN算法风格迁移验证；

2. GCN科学出版物分类导入依赖库、定义参数配置、转换数据格式为mindrecord；

四川美术学院基于CycleGAN算法风格迁移进行验证会议，总结问题

GCN科学出版物分类定义GCN网络参数和结构、定义训练和评估函数、启动训练和评估

完成智能基座课程GCN科学出版物分类案例markdown文档1个；

GCN科学出版物分类解决警告一个，替换为mindspore2.0接口;

完成四川美术学院基于CycleGAN算法风格迁移报告

1、GCN科学出版物分类解决警告一个，替换为mindspore2.0接口;

2、提交智能基座3个项目的PR,并解决push代码出现的问题，成功提交；

1. 利用download下载数据集在启智NPU运行自动问答实验；
2. 考试可信考试的软件工程与技术和信息安全考试；
3. 四川美术学院基于CycleGAN算法风格迁移验证；
4. GCN科学出版物分类导入依赖库、定义参数配置、转换数据格式为mindrecord；

进度：  
1、完成100%；  
2、完成100%；  
3、完成100%；  
4、完成100%；

1、四川美术学院基于CycleGAN算法风格迁移进行验证会议，总结问题；

2、GCN出版物分类定义GCN网络参数和结构、定义训练和评估函数、启动训练和评估；

3、完成智能基座课程GCN科学出版物分类案例markdown文档1个；

进度：  
1、完成100%；  
2、完成100%；  
3、完成100%；

1. 完成技术干货2篇;
2. 完成智能基座项目前沿网络案例-YOLOV3文档书写

2、安装MindStudio，并在mindspore+cpu+MindStudio环境下跑通基于LSTM的情感分类；

1、GCN科学出版物分类解决警告一个，替换为mindspore2.0接口;

2、提交智能基座3个项目的PR,并解决push代码出现的问题，成功提交；

3、完成技术干货2篇;

1. 完成智能基座项目前沿网络案例-YOLOV3文档书写；

进度：  
1、完成100%；  
2、完成100%；  
3、完成100%；  
4、完成100%；

元元总，这周工作内容

近期工作内容：

智能基座项目三个，已完成基于LSTM的情感分类和基于BERT的自动问答，已完成100%，GCN科学出版物分类正在开发中，已完成80%；

AI+x验收三个，成都中医院大学数字中医药分析项目和西南交大线稿提取项目已验收完毕，验收报告已写, 已完成100%；四川美术学院基于 CycleGAN 算法风格迁移正在验证中, 已完成50%。

1. 安装MindStudio，并在mindspore+cpu+MindStudio环境下跑通基于LSTM的情感分类；

2、和成都中医药接口人沟通问题，发送etrans上传数据集；

1、学习Attention机制和Transformer结构；

2、完成和修改高校项目验收报告2个；

3、和四川美院接口人沟通问题，预估项目代码完工时间；

1、和四川美院接口人沟通问题，预估项目代码完工时间；

2、学习Attention机制和Transformer结构；

1、完成和修改高校项目验收报告2个；

2、学习BERT和GPT模型；

1、和四川美院接口人沟通问题，预估项目代码完工时间；

2、学习Attention机制和Transformer结构；

3、完成和修改高校项目验收报告2个；

4、学习BERT和GPT模型；

进度：  
1、完成100%；  
2、完成100%；  
3、完成100%；  
4、完成100%；

1、修改完善《昇思MindSpore创新训练营-京津冀-课前准备》文档；

2、田奇介绍如何从基础大模型炼成行业大模型，总结与mindspore相关内容；

1. 验证四川美院算法风格迁移项目代码；
2. 华为专家开会评审AI+x两个项目；

2、和四川美院项目接口人反馈意见和交流；

1、修改完善《昇思MindSpore创新训练营-京津冀-课前准备》文档；

2、田奇介绍如何从基础大模型炼成行业大模型，总结与mindspore相关内容；

1. 验证四川美院算法风格迁移项目代码；
2. 华为领导专家开会评审AI+x两个项目，给出评审意见；

进度：  
1、完成100%；  
2、完成100%；  
3、完成100%；  
4、完成80%；

1. 完成四川美院算法风格迁移项目验证报告；

2、参与互联网+大赛相关会议2个，为互联网+做支撑准备；

实践

1、审核和复现四川美院算法风格迁移项目图文案例；

1. 完成技术干货2篇；

1、完成四川美院算法风格迁移项目验证报告；

2、参与互联网+大赛相关会议2个，为互联网+做支撑准备；

3、审核和复现四川美院算法风格迁移项目图文案例；

4、完成技术干货2篇；

进度：  
1、完成100%；  
2、完成100%；  
3、完成100%；  
4、完成100%；

1、完成转正答辩PPT，填写转正相关资料；

2、完成兰州大学机器学习原理mindspore示例代码编写2个；

3、完成兰州大学机器学习原理mindspore示例代码编写5个；

4、完成创新训练营需要准备前置课程测验题目5个；

进度：  
1、完成100%；  
2、完成100%；  
3、完成100%；  
4、完成100%；

1、完成转正答辩PPT，填写转正相关资料；

2、完成兰州大学机器学习原理mindspore示例代码编写2个；

1、完成兰州大学机器学习原理mindspore示例代码编写5个；

2、完成创新训练营需要准备前置课程测验题目5个；

1、学习GPT,为内部公开课讲课做准备；

1、和四川美院接口人沟通，审核验收图文案例；

2、完成兰州大学机器学习原理mindspore示例代码编写4个；

1、学习GPT,为内部公开课讲课做准备；

2、参加《机器学习原理》教材添加MindSpore代码示例的对齐的会议和《Linux网络诊断工具mtr》会议培训；

1、和四川美院接口人沟通，审核验收图文案例；

2、完成兰州大学机器学习原理mindspore示例代码编写4个；

3、学习GPT,为内部公开课讲课做准备；

4、参加《机器学习原理》教材添加MindSpore代码示例的对齐的会议和《Linux网络诊断工具mtr》会议培训；

进度：  
1、完成100%；  
2、完成100%；  
3、完成100%；  
4、完成100%；

1、验收西南交大线稿提取项目图文案例；

2、验证[step\_into\_chatgpt](https://github.com/mindspore-courses/step_into_chatgpt)课件代码并添加图片，解决bug5个；

1、验证AI+X三个项目，进行技术评审会议；

2、书写AI+X二个项目技术评审会议纪要；

1、验收西南交大线稿提取项目图文案例；

2、验证[step\_into\_chatgpt](https://github.com/mindspore-courses/step_into_chatgpt)课件代码并添加图片，解决bug5个；

3、验证AI+X三个项目，进行技术评审会议；

4、书写AI+X二个项目技术评审会议纪要；

进度：  
1、完成100%；  
2、完成100%；  
3、完成100%；  
4、完成100%；

1、修改西南交大线稿提取项目图文案例；

2、进行ChatGLM2-6B推理，解决BUG一个，但还有BUG需要解决；

1. 参加兰州大学机器学习例会；
2. 准备并进行大模型技术公开课，讲解GPT2;

1、修改西南交大线稿提取项目图文案例；

2、进行ChatGLM2-6B推理，解决BUG一个，但还有BUG需要解决；

1. 参加兰州大学机器学习例会；
2. 准备并进行大模型技术公开课，讲解GPT2;

进度：  
1、完成100%；  
2、完成100%；  
3、完成100%；  
4、完成100%；

1、准备并进行大模型技术公开课，讲解prompt;

2、验证新版西南交大线稿提取项目图文案例，跑通代码；

1. 完成ChatGLM2命令行交互，可以进行对话；

2、消除ChatGLM2推理和命令行交互对话所产生的warnings,并提交PR;

刘卫星本月1-17号的产出：  
1.win+cpu上安装mindspore  
2、在启智平台上运行Sentiment情感分析实验  
3.基于BERT的自动问答实验，通过关闭xGate后，成功将数据集传送到启智平台上  
4.基于BERT的自动问答实验，导入依赖库、运行环境配置和定义超参数   
5. 数据预览、定义数据预处理函数和定义训练函数  
6.中药炮制饮片质量判断模型准确性验证   
7. 可信认证考试，含编程知识与应用、软件工程与技术  
8. 润和软件可信认证编程能力题板块考试   
9. 西南交大经图项目验证，修复bug3个，还有bug需要处理

刘卫星从5月18开始到今天的月度产：

1、 基于BERT的自动问答实验，执行训练、定义评估函数、评估模型效果和单例评估；

2、书写基于BERT的自动问答实验markdown文档；

3、在启智NPU环境上运行成都中医院大学数字中医药分析项目；

4、书写成都中医院大学数字中医药分析项目验收报告；

5、西南交大线稿提取项目在CPU上训练、测试和交互测试；

6、基于BERT的自动问答实验，消除5个WARNING；

7、西南交大线稿提取项目 CycleGan\_DexiNed启智NPU上训练、测试和交互测试；

8、召集会议评审西南交大线稿提取项目；

9、西南交大线稿提取项目启智NPU上910推理跑通；

10、西南交大线稿提取项目验收报告书写；

11、和华为员工沟通自动问答实验数据集上传并返回URL链接；

12、学习软件工程与技术和数据结构与算法，为可信考试做准备；

13、利用download下载数据集在启智NPU运行自动问答实验；

14、考试可信考试的软件工程与技术和信息安全考试；

15、 四川美术学院基于CycleGAN算法风格迁移验证；

16、GCN科学出版物分类导入依赖库、定义参数配置、转换数据格式为mindrecord；

17、四川美术学院基于CycleGAN算法风格迁移进行验证会议，总结问题；

18、GCN出版物分类定义GCN网络参数和结构、定义训练和评估函数、启动训练和评估；

19、完成智能基座课程GCN科学出版物分类案例markdown文档1个；

20、GCN科学出版物分类解决警告一个，替换为mindspore2.0接口;

21、提交智能基座3个项目的PR,并解决push代码出现的问题，成功提交；

工作总结：

1. 环境安装，win+cpu上安装mindspore，安装好Anaconda、pycharm，学习启智NPU调试任务创立和使用；
2. 在启智平台和MindStudio上运行Sentiment情感分析实验，本实验以IMDB影评情感分类为例来体验MindSpore在自然语言处理上的应用，使用MindSpore框架进行自然语言中处理情感分类问题，理解了如何通过定义和初始化基于LSTM的SentimentNet网络进行训练模型及验证正确率，同时也熟悉了MindSpore深度学习框架的基础用法，以及MindStudio软件的简单操作等。
3. 基于BERT的自动问答实验，步骤如下：导入依赖库、运行环境配置和定义超参数、数据预览、定义数据预处理函数和定义训练函数、执行训练、定义评估函数、评估模型效果和单例评估。基于MindSpore实现阅读理解问答系统，使用了BERT预训练微调的方式进行训练。
4. GCN科学出版物分类，步骤如下：导入依赖库、定义参数配置、转换数据格式为mindrecord、定义GCN网络参数和结构、定义训练和评估函数、启动训练和评估。使用MindSpore实现图卷积神经网络GCN，用于科学出版物分类。
5. AI+X三个项目验证，成都中医院大学数字中医药分析项目、西南交大线稿提取项目和四川美术学院基于CycleGAN算法风格迁移。

工作中的问题及思考：

1. 必须关闭xGate后，才能成功将数据集传送到启智平台上；

2、基于BERT的自动问答实验，保证保存和加载时的网络结构必须一致，不然会出现警告；

3、提交智能基座3个项目的PR, push代码出现port 443: Bad access问题，通过使用 token，来解决这个问题，成功提交；

四川美术学院基于CycleGAN算法风格迁移已经按照反馈意见进行修改，图文案例按照readme的格式修改了；

成都中医院大学数字中医药分析项目、西南交大线稿提取项目还在修改中，目前没有反馈。

元元总，由于工资不高和想回长沙发展，我想离职，是不是交接完工作就行了？还是走怎样的流程？

春爱姐，由于工资不高和想回长沙发展，我想离职，是不是交接完工作就行了？还是走怎样的流程？

下一步工作计划：

1. 将未完成的四川美院算法风格迁移项目继续验证；
2. 作为昇思MindSpore创新训练营助教，回答学生问题；
3. 不断学习，学习NLP、CV、大模型等，提高代码能力；

刘卫星7月产出：

1、学习Attention机制和Transformer结构；学习BERT和GPT模型；

2、完成和修改高校项目验收报告2个；

3、修改完善《昇思MindSpore创新训练营-京津冀-课前准备》文档；

4、 验证四川美院算法风格迁移项目代码；

5、 华为专家开会评审AI+x两个项目；

6、参与互联网+大赛相关会议2个，为互联网+做支撑准备；

7、审核和复现四川美院算法风格迁移项目图文案例；

8、 完成技术干货2篇；

2、完成兰州大学机器学习原理mindspore示例代码编写7个；

10、完成创新训练营需要准备前置课程测验题目5个；

并且有相应的对抗鉴别器 D y和 Dx。

可视化展示3行10列，随机展示

@所有人 各位老师好，AI+X西南交大线稿提取项目需要各位评审，想定在下周一上午9：30到11点的时间进行，不知各位是否有空？

https://github.com/mindspore-courses/step\_into\_llm/pull/4

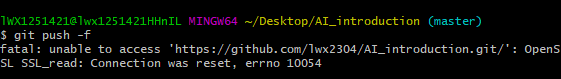
@吕昱峰 昱峰老师，这是ChatGLM2命令行交互的PR，通过终端设置export GLOG\_v=3才可以消除warnings。Python脚本也都是在终端运行的。



Etrans

婉怡姐，这个是自然语言处理图神经网络GCN科学出版物分类实验的代码包和数据包，辛苦帮忙上传一下，给个url.

@苏婉怡@罗琳琳，两位老师，这是自然语言处理基于BERT的自动问答实验的代码包，辛苦帮忙上传一下，给个url.

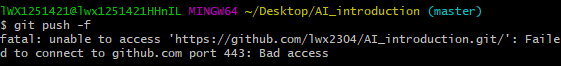


成都中医药进展：接口人询问我几个问题，将网络模型API接口（包括损失函数、优化器）写成.py的形式；将训练和推理脚本写成.sh形式；然后正在继续书写中；

西南交大进展：西南交大接口人已完成所有交付件，图文案例详细清晰；

四川美院进展：以前的反馈意见反馈给了接口人，他们正在攻关中，目前还没有提交东西给我；

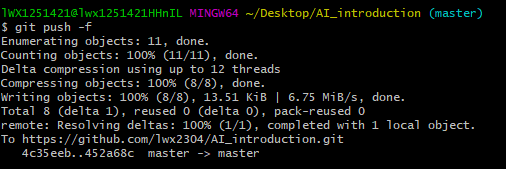
提交PR失败，网上总结的方法



通过

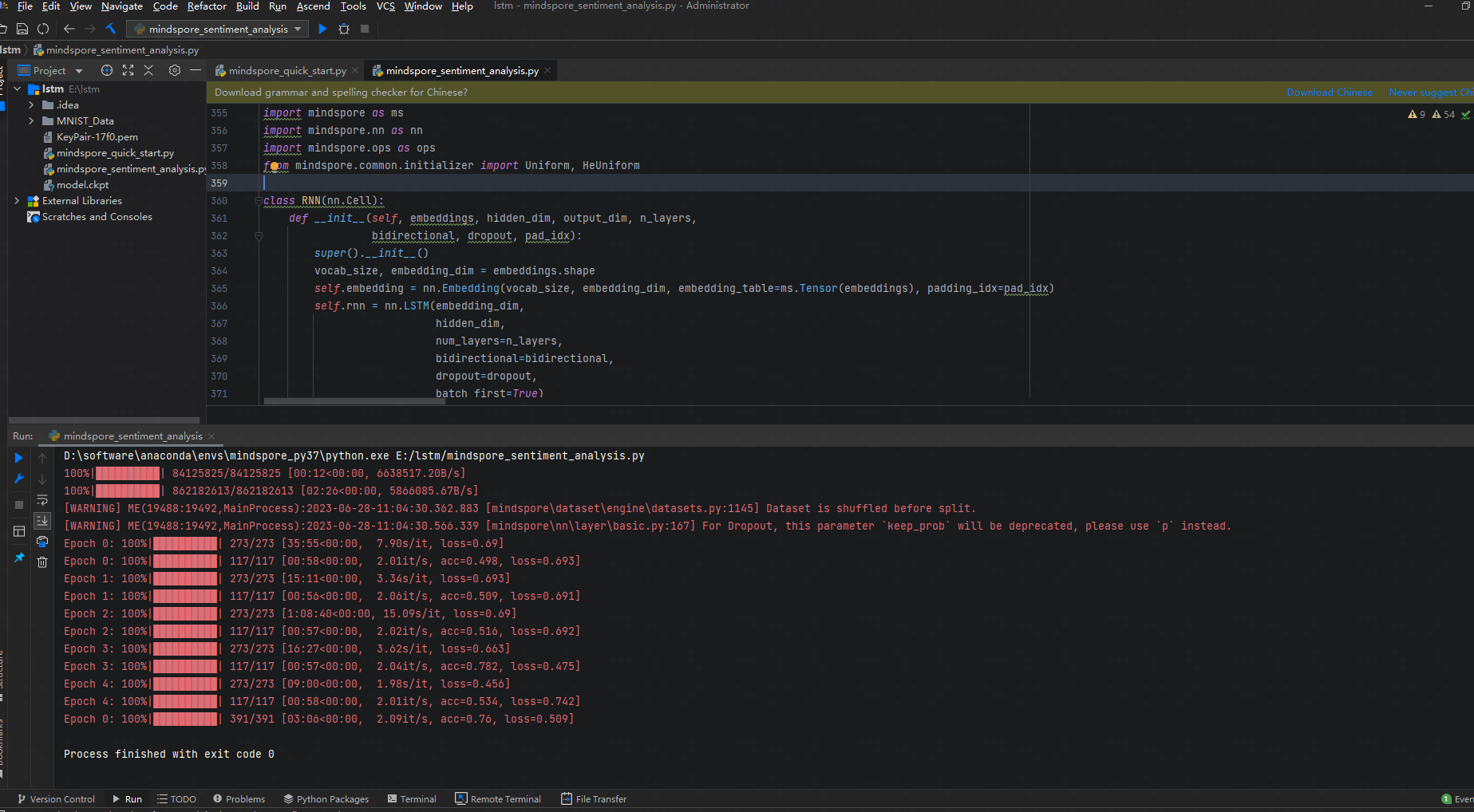
<https://blog.csdn.net/weixin_45844049/article/details/123733065>

来解决



成功

Github自己建仓



责任田描述：

1. 智能基座自然语言处理项目3个；
2. AI+x西南交大、四川美院和成都中医药大学项目验收；

2023/5/18 ~ 2023/6/30

西南交大线稿提取项目验收周进展：交付件都通过。

成都中医药数字中医药分析项目验收周进展：文件格式为ipynb的端到端可执行的图文案例在下载数据时出错，因为URL链接是私有的，得重新上传数据集生成一个URL链接，然后运行。其余交付件没问题。

四川美院数字美术学分析项目验收周进展：已催促四川美院接口人抓紧时间，专家正在抓紧开发代码。

昱峰老师，成都中医药数字中医药分析项目图文案例在下载数据时出错，因为URL链接是私有的，得重新上传数据集生成一个URL链接。麻烦你上传数据集并返回给我一个URL。

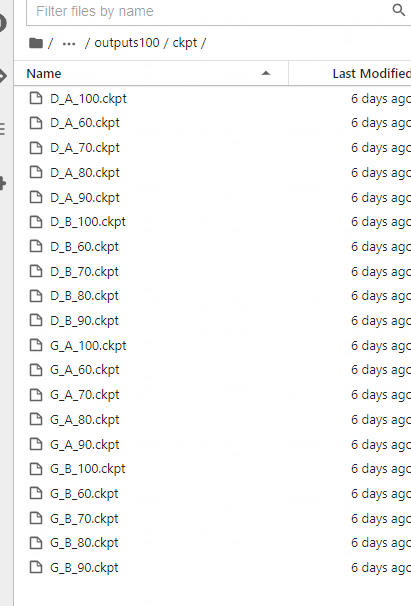




@苏婉怡 @罗琳琳，两位老师，这是AI+X成都中医药大学数字中医药分析项目的数据集，比较大有15G,辛苦帮忙上传一下，给个url.

<https://onebox.huawei.com/p/0cf9c687020ab118c7899e4f2f418104>

数据集展示做出3行4列，3行对应3种药品，4列对应4种状态。共12类。



改成60个epochs

bash scripts/run\_train\_ascend.sh ./data/dataset 100 1 300

bash scripts/run\_train\_ascend.sh ./data/dataset 60 1 300

import numpy as np

import mindspore

import mindspore.ops as ops

from mindspore import Tensor

def linear\_loss(y\_true, y\_pred):

"""

:param y\_true: 真实值

:param y\_pred: 预测值

"""

n = len(y\_true)

loss = np.sum(ops.abs(y\_true - y\_pred)) / n

# loss = (ops.abs(y\_true - y\_pred)).sum() / n

return loss

# 测试代码

# y\_true = Tensor(np.array([1, 2, 3, 4, 5]), mindspore.float32)

# y\_pred = Tensor(np.array([1.5, 2.5, 3.5, 4.5, 5]), mindspore.float32)

# print(f"linear\_loss: {linear\_loss(y\_true, y\_pred):.4f}")

if \_\_name\_\_ == "\_\_main\_\_":

from mindspore import Tensor

np.random.seed(11070109)

x = np.random.randn(3, 3)

#x = paddle.to\_tensor(x).cast("float32")

x = Tensor(x, mindspore.float32)

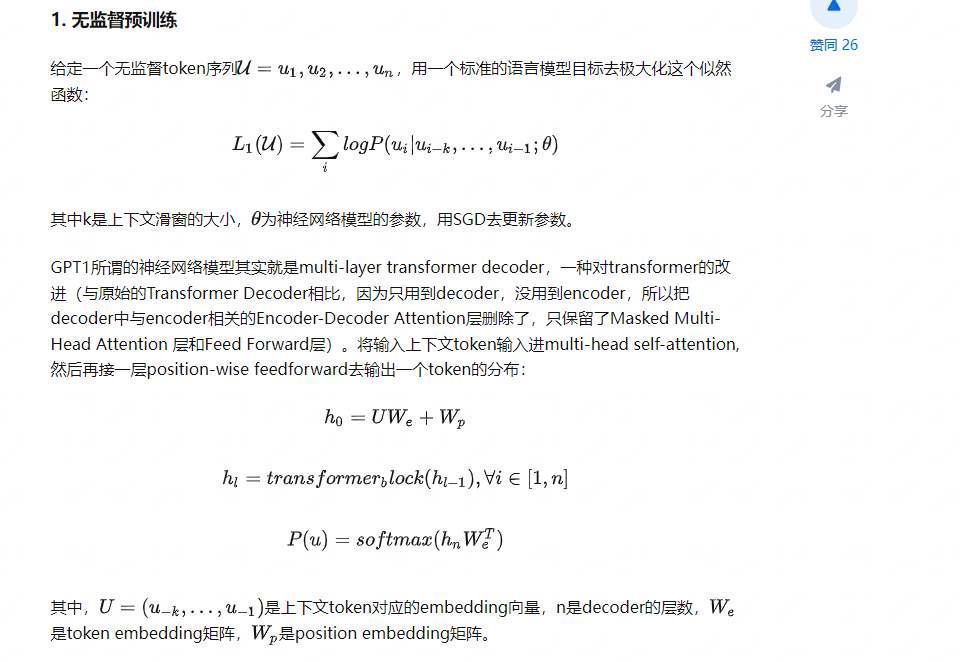
y = (x > 0.5).cast("float32")

loss = AsymmetricLossOptimized()

out = loss(x, y)

元元总，这周六我换租房的地方，约好搬家，因此周六就不加班哈

春爱姐，刚好周六我换租房的地方，得搬家，因此周六就不加班了，我也对元元总说了。



import math

ALPHA = 0.3

DIFF = 0.00001

def predict(theta, data):

results = []

for i in range(0, data.\_\_len\_\_()):

temp = 0

for j in range(1, theta.\_\_len\_\_()):

temp += theta[j] \* data[i][j - 1]

temp = 1 / (1 + math.e \*\* (-1 \* (temp + theta[0])))

results.append(temp)

return results

def training(training\_data):

size = training\_data.\_\_len\_\_()

dimension = training\_data[0].\_\_len\_\_()

hxs = []

theta = []

for i in range(0, dimension):

theta.append(1)

initial = 0

for i in range(0, size):

hx = theta[0]

for j in range(1, dimension):

hx += theta[j] \* training\_data[i][j]

hx = 1 / (1 + math.e \*\* (-1 \* hx))

hxs.append(hx)

initial += (-1 \* (training\_data[i][0] \* math.log(hx) + (1 - training\_data[i][0]) \* math.log(1 - hx)))

initial /= size

iteration = initial

initial = 0

counts = 1

while abs(iteration - initial) > DIFF:

print("第", counts, "次迭代, diff=", abs(iteration - initial))

initial = iteration

gap = 0

for j in range(0, size):

gap += (hxs[j] - training\_data[j][0])

theta[0] = theta[0] - ALPHA \* gap / size

for i in range(1, dimension):

gap = 0

for j in range(0, size):

gap += (hxs[j] - training\_data[j][0]) \* training\_data[j][i]

theta[i] = theta[i] - ALPHA \* gap / size

for m in range(0, size):

hx = theta[0]

for j in range(1, dimension):

hx += theta[j] \* training\_data[i][j]

hx = 1 / (1 + math.e \*\* (-1 \* hx))

hxs[i] = hx

iteration += -1 \* (training\_data[i][0] \* math.log(hx) + (1 - training\_data[i][0]) \* math.log(1 - hx))

iteration /= size

counts += 1

print('training done,theta=', theta)

return theta

if \_\_name\_\_ == '\_\_main\_\_':

training\_data = [[1, 1, 1, 1, 0, 0], [1, 1, 0, 1, 0, 0], [1, 0, 1, 0, 0, 0], [0, 0, 0, 0, 1, 1], [0, 1, 0, 0, 0, 1],[0, 0, 0, 0, 1, 1]]

test\_data = [[0, 1, 0, 0, 0], [0, 0, 0, 0, 1]]

theta = training(training\_data)

res = predict(theta, test\_data)

print(res)







中医药反馈意见：

1. markdown中图片显示不出来；
2. 推理部分显示的图片需要备注，中英文对照；

MindSpore2.0.0a0\_cann\_6.0.rc1.alpha005

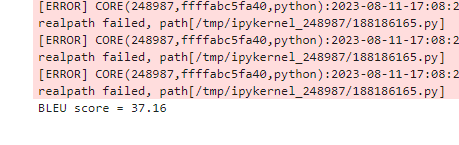
1.Transformer有一个BUG没解决；6.CodeGeeX和7.Prompt你说的先不验证

其余的都验证完了

庹元元；吕昱锋；赵婷；谭志鹏

赵宇老师，这是验证后的代码。其中Transformer代码中的数据集下载url替换为译壬提供的最新的，可以跑通。









青林，你好。请问这个bug怎么解决？

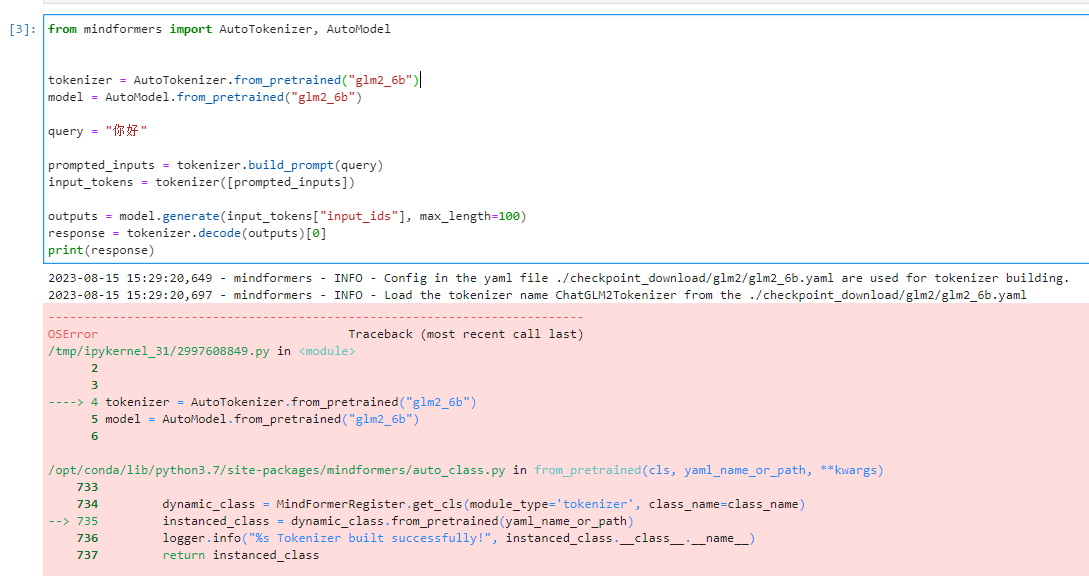
这里mindformers安装采用的是源码编译安装，环境是启智NPU+mindspore2.0

git clone -b dev https://gitee.com/mindspore/mindformers.git

cd mindformers

bash build.sh

青林，在启智GPU+mindspore2.0环境下第一个bug解决了，但又出现这样的bug，请问如何解决？





GPT2:

Bert抢了风头，transformer encoder 预训练，各种刷榜，显得GPT不是那么重要的工作

1. 对GPT进行小的改动，改动不多，还是transformer decoder 结构。调整Transformer：\*\*将layer normalization放到每个sub-block之前，并在最后一个self-attention后再增加一个layer normalization，对结果进行平均或者平滑，从而达到更好的收敛效果。

layer normalization前置比layer normalization后置效果要好一些，有专门的论文讨论。

2、把规模扩大，把模型参数从1亿扩大到1.5B，也就是15亿。

Language model 本身并不复杂，可以用大规模的数据去训练，从而学习到很多信息，学到各种知识。

输入法，前面的字预测后面的字，是一个小型的Language model。

输入前n-1个词，预测下一个词，取argmax(-1)即可得到预测的最好的词。

GPT2 transformer decoder 结构，自回归式的预测

A robot may not

预测的过程是预测出一个词后，把这个词作为输入再预测下一个词，如此循环下去，知道预测出SEP则结束。

预训练时做过多任务instruction的训练

预训练的数据里面做过不同指令的多任务的训练，后面加上新的数据，数据没有标签，可以做到zero-shot learning

GPT2任务设定和GPT1不同，想做小样本甚至零样本的学习，做多种任务的训练，

训练好一个模型后，能够不给定标签，迁移到另一个任务。

zero-shot learning**零样本学习就是让学习器对其**从来没有见过的类别**进行分类**。比如：给学习器**一堆马和老虎的图片进行训练，训练完毕后，我们输入一张斑马的照片**，并希望学习器告诉我**们“这是斑马！**”。

跟人类思维方式接近，人类可以举一反三。

百亿级别就可以有涌现了

大模型千亿级别的得配上千亿级别的数据，如果数据不够多，模型训练会出现欠拟合现象

就好像背单词，当你背会5000个单词也许没什么用，但当你背会1万个，效果一下子就有了

这种突破性进展，是知识的长期积累涌现而来。

当一个人的知识积累到一定程度，就会出现涌现效应。

知识越多，涌现效应越明显。

我们每一次学习新的知识，大脑中就会产生更多的信息链接。

链接越多，可能性就越多，大脑就越具有创造性。

把一件事坚持做下去，必会迎来质变的那一刻。

GPT模型主要包含两个阶段，第一个阶段，先利用大量未标注的语料预训练一个[语言模型](https://so.csdn.net/so/search?q=%E8%AF%AD%E8%A8%80%E6%A8%A1%E5%9E%8B&spm=1001.2101.3001.7020)，接着，在第二个阶段对预训练好的语言模型进行微改，将其迁移到各种有监督的NLP任务，并对参数进行fine-tuning。

GPT1、Bert大概1亿参数，12层，512 context window

GPT2拉到48层, 1024 context window

context window 大小代表能传入的输入token的数量，这个很重要。

在昇腾上面加混合精度，显存不够大，所以进行梯度累加

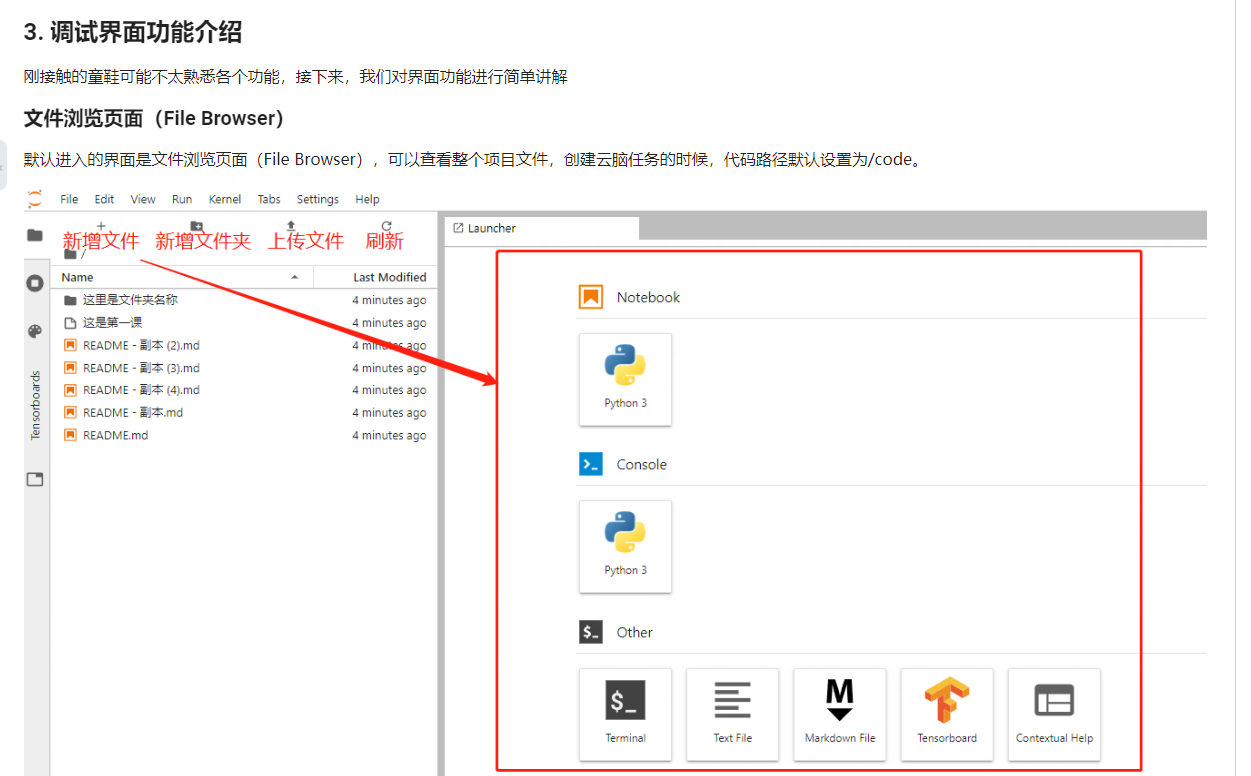
加混合精度是用float16计算的，float16表示的数字范围比float32小很多

is\_finite = all\_finite(grads, status)

if is\_finite:

True 没有溢出

False 溢出





* 方式1：源码编译安装

支持源码编译安装，用户可以执行下述的命令进行包的安装

git clone -b dev https://gitee.com/mindspore/mindformers.git

cd mindformers

bash build.sh

Pretrain+finetune 相比prompt不足之处：

1. 当模型参数越来越大，finetune不动
2. Pretrain+finetune适配下游任务，对于模型部署不太友好。

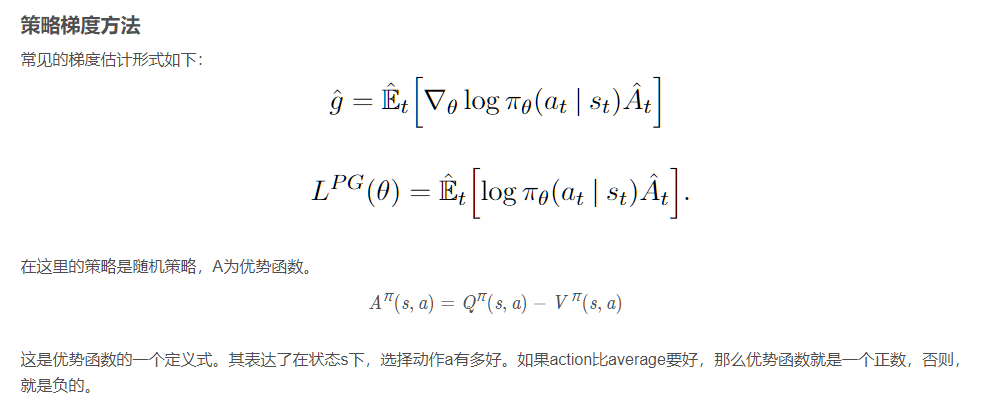
Pretrain 学习大量知识，prompt通过提示调用特定的部分知识来推理和预测

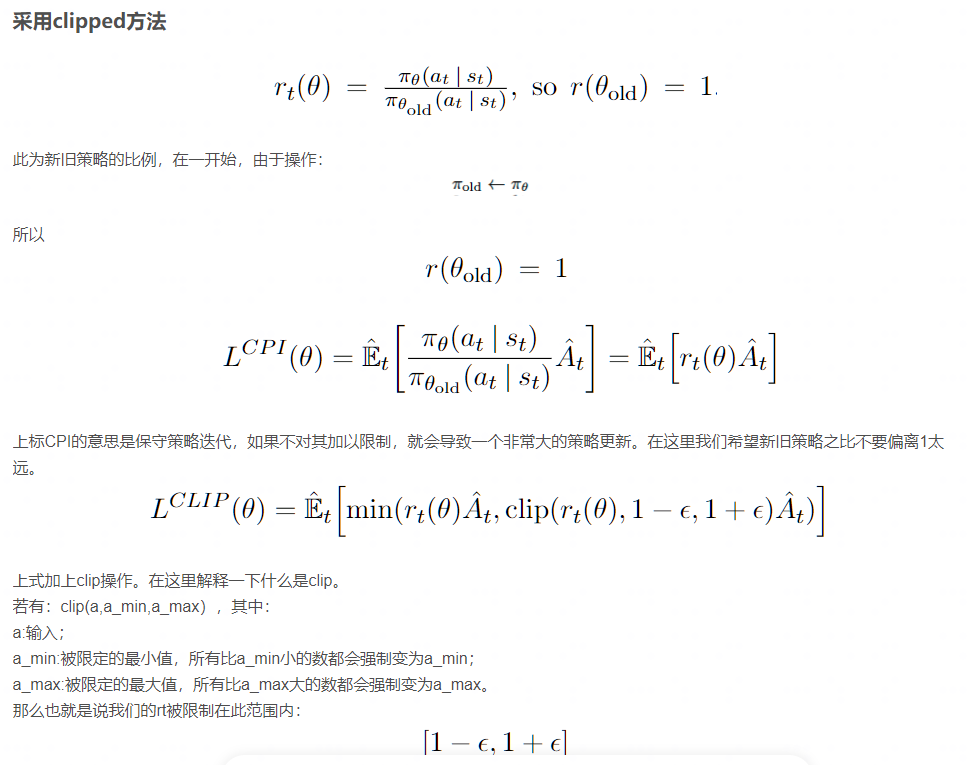
prompt提示或者通过完形填空的方式预测所要的结果

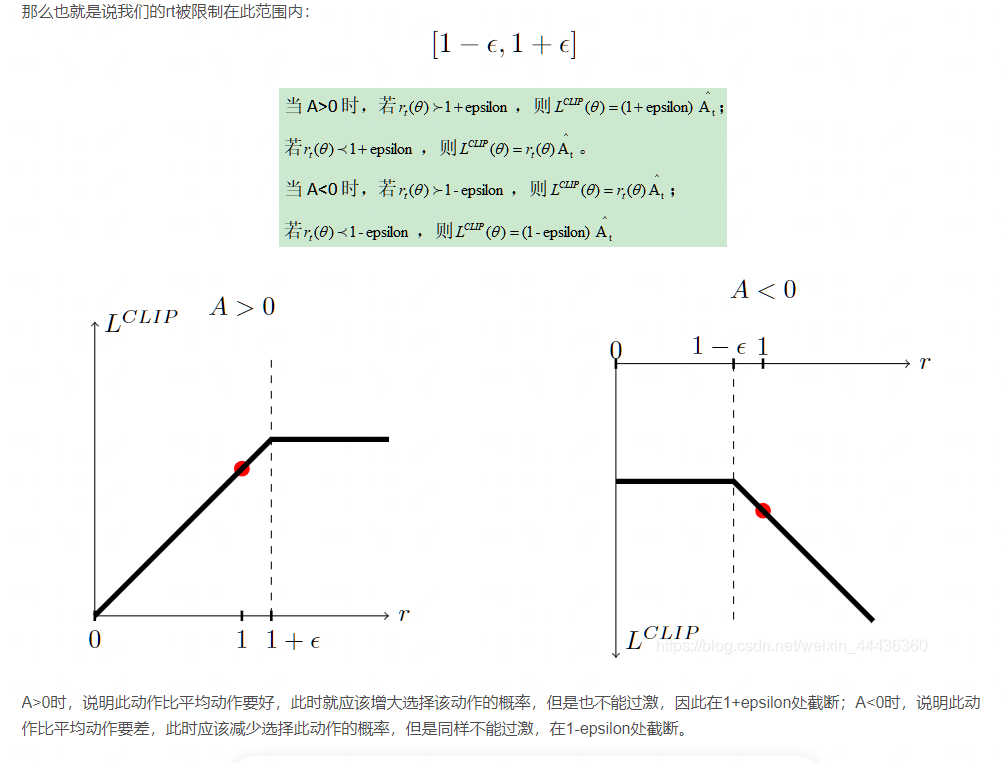
Prompting 中最主要的两个部分为 template 与 verbalizer 的设计。

他们可以分别基于任务类型和预训练模型选择(shape)或生成方式(huamn effort)进行分类。

python chatglm2.py 作者：拉黑小天才 https://www.bilibili.com/read/cv24121178 出处：bilibili







强化学习像是教小孩做事一样，小孩做好事的话给他奖励，做坏事给他惩罚，这样小孩就养成一个好的习惯。

## 交互测试结果示例图



图4：交互测试预测调整图片



图5：交互测试预测结果图片