**AI期中project报告**

24212020231 叶天勤

**运行方式：**

进入目录下后，先设置环境并安装必要的库，如：

pip install transformers datasets torch

运行训练模型

python train\_model.py

运行训练模型的时间较长，我上传文件当中应该已经将我这边生成好的模型一并上传了，这一步便可以省略。

最后运行生成语句文件

python generate\_sentence.py

**使用模型：**

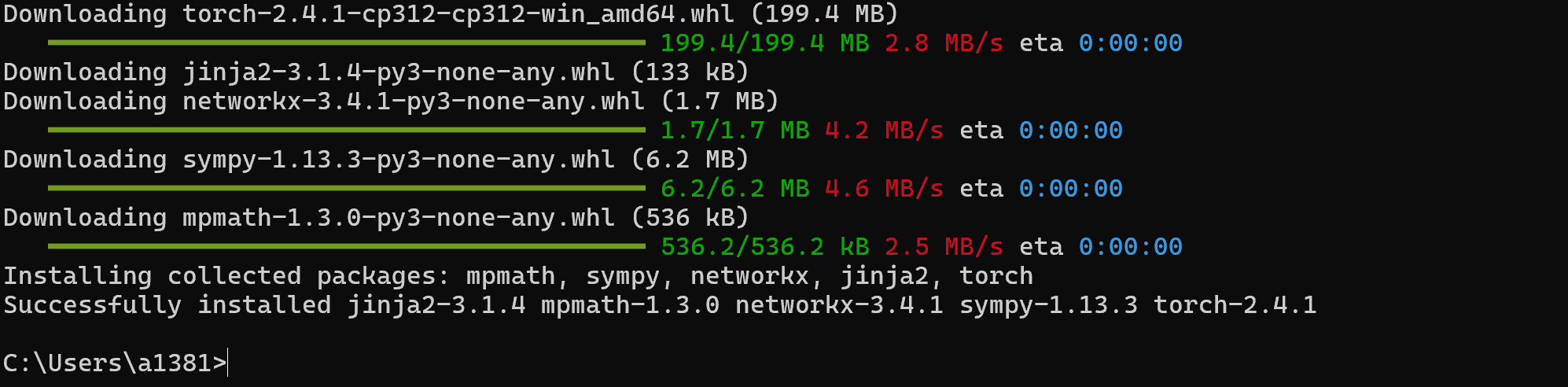
模型：GPT-2

数据集：wikitext-2-raw-v1

**具体介绍：**

首先在选择模型当中，对于一个入门项目,所以我选择一个较小的模型,最后选定了GPT-2。而使用 Hugging Face 的 wikitext-2-raw-v1 数据集可以让我直接从 Hugging Face 加载数据，而不需要手动下载和处理文件。便于了代码的编写。

接着是安装必要的库，如下图所示，这一步就不再赘述。



编写好训练模型代码以后，开始运行训练代码。解决了一些语法上的错误，正确开始运行时，可以发现模型训练所需要的时间特别长（最后实际花费了18个小时）。我进行了这方面的研究，发现我使用的是CPU进行的运行，而如果使用GPU可以大大进行加速。我尝试设置使用GPU进行运行，而进一步研究后发现NVIDIA的GPU可以更为方便地在代码中进行设置，而我是AMD的GPU，因而只能暂且放弃。下图是运行训练模型完成后的截图。



完成模型训练后，在代码中训练的模型位置处于./fine\_tuned\_gpt2当中，因而后续的生成句子代码需要锁定到这一目录。

完成生成句子的相关代码编写后，我最主要遇到的障碍是：尝试重塑（reshape）一个空张量时出现了问题。为了解决这个问题，要确保我的输入不是空的，所以我在生成时添加一个起始提示（prompt），我这里设置的是Once upon a time，所以可以看到后续生成的句子开头都有Once upon a time。不知道还有没有更好的解决办法。

最后运行结果如图所示。

