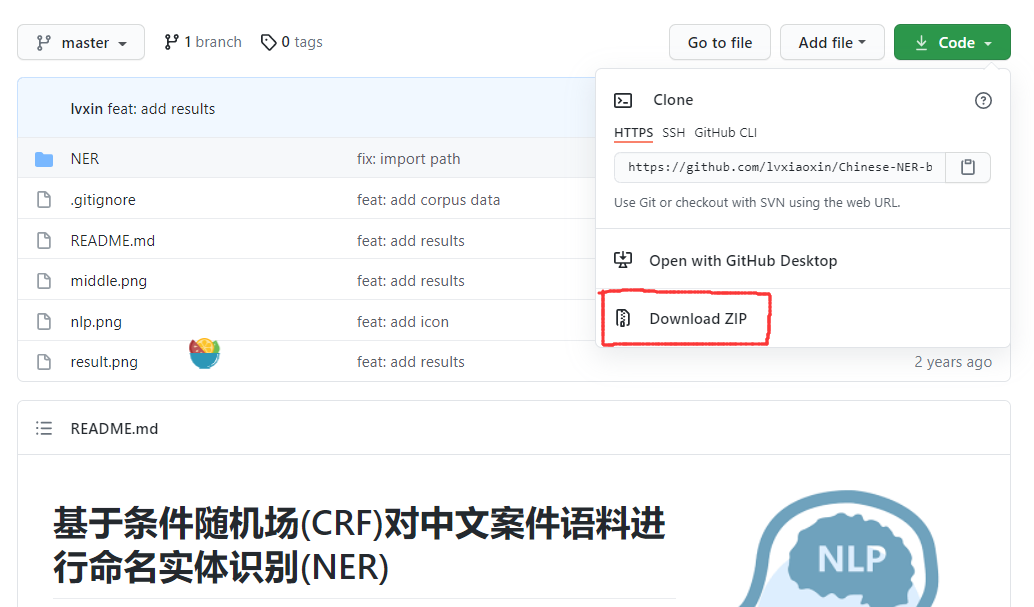
**实验步骤：**

1. 从GitHub中下载代码，GitHub地址：

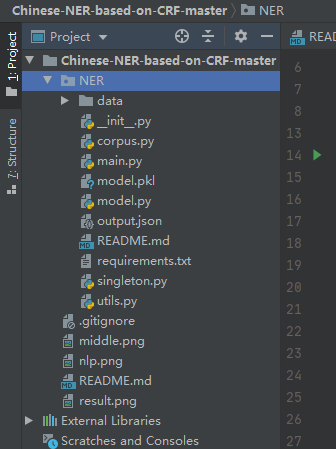
<https://github.com/guetlyy/Chinese-NER-based-on-CRF>

点击绿色code下载按钮，选择Download ZIP即可下载。



1. 下载完成后将会得到一个压缩包对压缩包进行解压，可以看到解压后的目录结构。
2. 使用你的开发环境打开步骤2解压的文件夹。

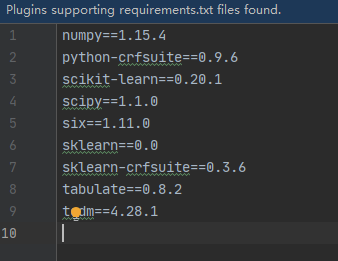
根据个人情况配置好此工程的Python解释器等其他设置。



1. 安装此工程运行所需要的依赖包。

在文件目录中的可以看到有requirements.txt的文件，请打开此文件，查看依赖包有哪些，你的环境是否已经安装有部分要求的依赖包，如果有请删除此文件中要求安装的依赖包，节省安装的时间。运行命令如下：

pip install -i [https://pypi.tuna.tsinghua.edu.cn/simple -r requirements.txt](https://pypi.tuna.tsinghua.edu.cn/simple%20-r%20requirements.txt)



安装依赖包时，你可能会遇到此错误提醒：

请切换到requirements.txt所在的目录下，再运行安装依赖包的命令。

在pycharm中的Terminal下安装依赖包。

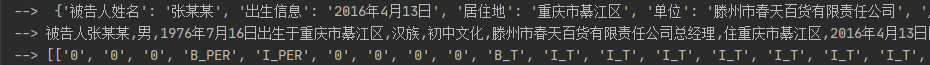
可能用到的命令：

激活虚拟环境：activate your\_env\_name

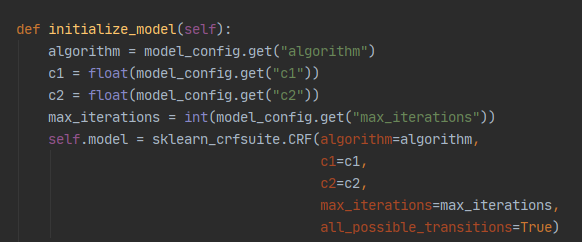
切换不同逻辑盘符(如切换到E盘)：e:

同逻辑盘切换目录：cd your\_dir

1. 安装好依赖包后，运行main.py文件，可以看到目录中增加了model.pkl和output.json文件，前者是训练后保存的模型，后者是测试集识别的结果输出到文件，打开查看实体识别结果。根据课堂所学的BIO实体标注，查看打印输出的标注结果是否正确，标注错误的地方请指出并记录下来。



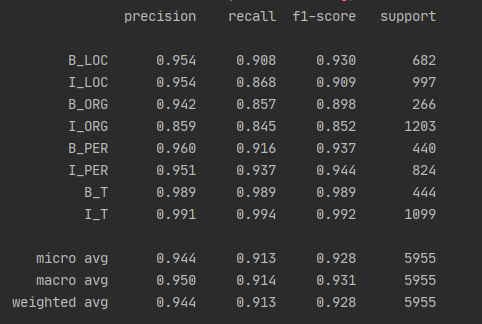
1. 打开model.py文件，查看调用随机场的函数的代码



打开utils.py 文件，查看使用的随机场的参数，c1为L1范数正则化系数，c2为L2范数正则化系数。



修改参数，再重新运行main.py，观察输出的标注准确率。



1. 更改程序为预测模式，打开utils.py，找到训练的配置参数\_is\_train，将参数由True改为False，在main.py文件末尾填写预测程序并给一小段文字进行预测。模型加载代码已给出，完成调用预测的函数即可，仿造main.py的36到42行代码，记录测试结果。