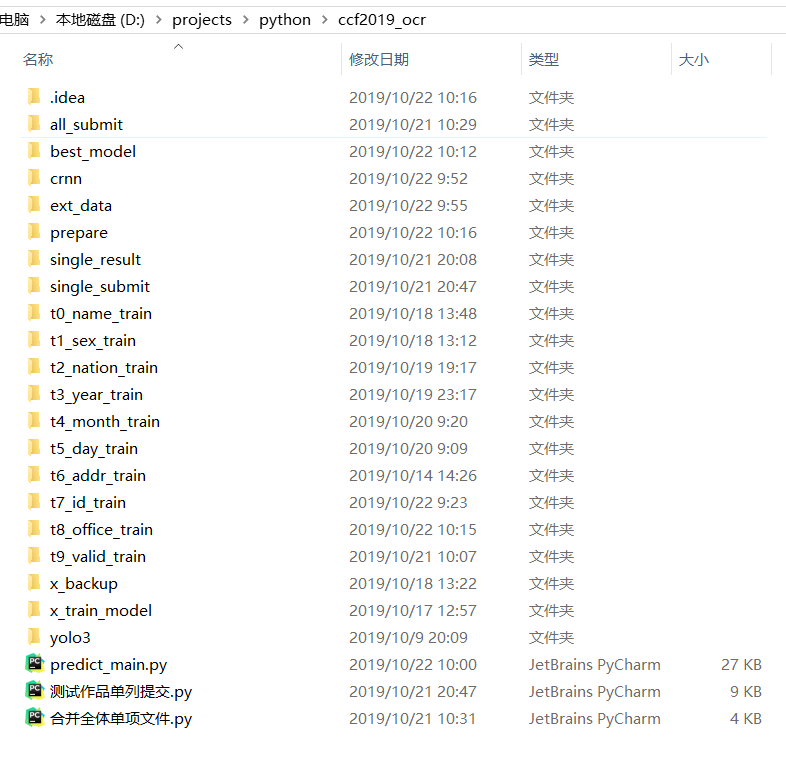
Python 3.7

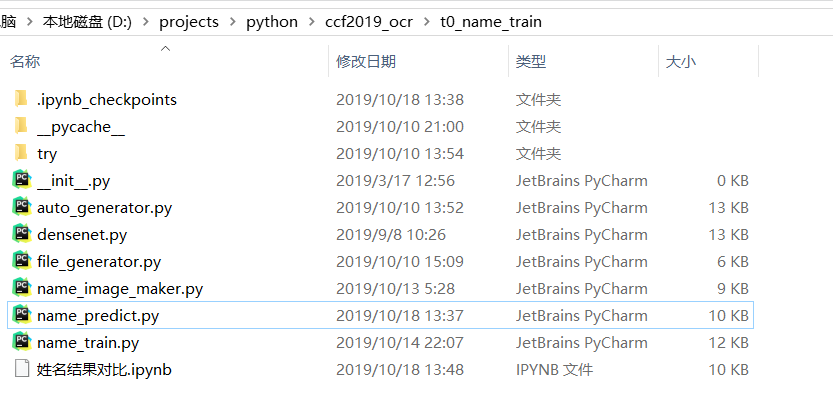
文件夹结构：



文件夹说明：

|  |  |
| --- | --- |
| All\_submit | DF提交的全体字段文件夹 |
| Single\_result | 单个模型预测的结果，没有进行提交整理（其余字段补空）的结果集，文件由：各自模型生成结果 |
| Single\_submit | 可以进行单个提交的文件夹，已经没有预测的字段补空，文件由：测试作品单列提交.py 生成 |
| Predict\_main.py | 根据yolo生成坐标文件，生成的文件仅是为了我单独测试方便，整合时文件应该不输出。直接交给后续模型处理 |
| yolo3 | 显维提供 |
| crnn | 显维提供 |
| x\_train\_model | 单独训练的模型时临时存放权重文件。 |
| x\_backup | 部分代码备份，可以忽略 |
| ext\_data | 外部来源数据存放地点 |
| t0\_name\_train | 姓名 训练代码 |
| ~ |  |
| t9\_valid\_train | 有效期 训练代码 |

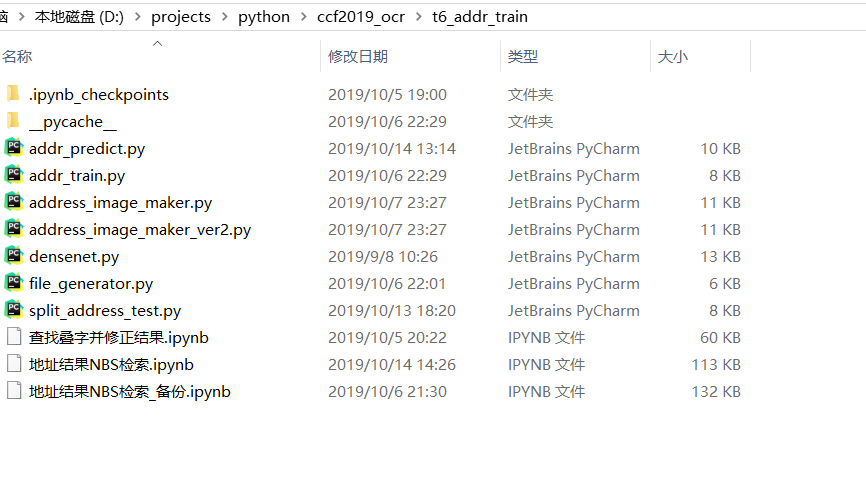
姓名训练模块说明：



|  |  |
| --- | --- |
| Name\_train.py | 训练代码 |
| Name\_predict.py | 预测代码 |
| auto\_generator.py | 废弃，训练时自动生成数据，太慢 |
| densenet.py | 显维提供，在crnn里已经存在，可以忽略 |
| name\_image\_maker.py | 训练数据多线程生成程序，用于生成训练数据。 |
| 姓名结果对比.ipynb | 结果之间的对比，方便辨别问题的原因。 |

※各个模块都是类似的结构，所以不重复说明

地址 训练模块：



|  |  |
| --- | --- |
| addr\_train.py | 训练代码 |
| addr\_predict.py | 预测代码 |
| file\_generator.py | 训练时，对数据做预处理。预测时 不用 |
| split\_address\_test.py | 因为地址是多行数据，进行按行二次切分。 |
| densenet.py | 显维提供，在crnn里已经存在，可以忽略 |
| address\_image\_maker.py | 训练数据多线程生成程序，用于生成训练数据。 |
| address\_image\_maker\_ver2.py | 第二次适当修改了数据，在第一的基础上继续生成数据，数据是累加的。 |
| 查找叠字并修正结果.ipynb | 废弃 |
| 地址结果NBS检索.ipynb | 对预测结果，根据国家统计局数据进行纠正修复。  如果需要整合，是不是可以放在最后一次性处理？ 不用放到模型预测里？ |