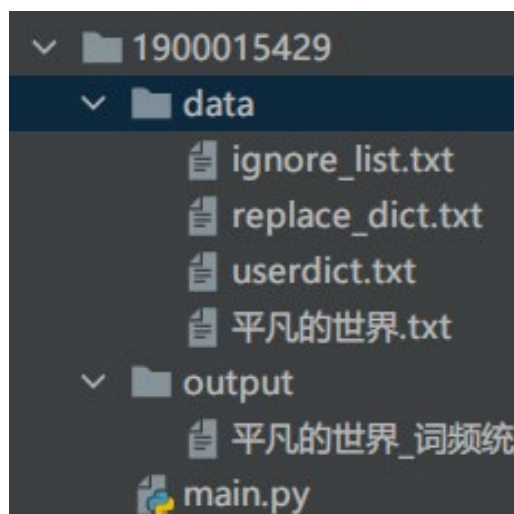


本程序主要实现了对《平凡的世界》进行分词和词频统计的功能，依赖的库是 jieba。

选择《平凡的世界》主要原因是，这本书的主要人物姓名较好辨认，并且主人公会以不同的名字出现在作品中，（如孙少平在文中也会以“少平”的名字出现）方便实现词语的合并



统计功能。

以上是代码文件的结构，其中 ignore\_list.txt 和 replace\_list.txt 分别提供了忽略词和替换词的字典，当然，这二者也可以在程序内部重新指定。

本程序最大的亮点是将整个分词和词频统计的过程封装到 TextAnalysis 类中，方便了代码的复用，关于该类的属性和方法见下图（关于属性和方法的详细内容见文件注释）：

