项目说明：

代码：

simple gensim lda包含gensim lda一个简单的例子

gensim lda包含使用gensim lda对20newsgroup分类的例子

lda with more dataset包含使用gensim lda对路透社新闻和imdb电影评论分类的例子

1 py文件正文中的代码实现主要功能，注释中的代码为组件，需要将组件插入正文中。

2 lda模型生成的各个主题的顺序是随机的，同时对概率几乎为0的主题会省略，所以得到主题的结果后，需要手动找出各个主题及分类逻辑。

3 代码中包含注释说明。

数据：

数据集包含在python库中，可引用库得到