都用gensim包里来做就行，不需要ubuntu

1. 先运行preprocess.py，得到字典和语料库
2. 修改主题数目。打开batch\_vs\_online.py, 先修改参数。chunksize=40, passes=50。修改batch\_vs\_online(新文件名)
3. 结果不好的话，是否需要修改构造单词云的策略？
4. 运行12分钟