现在能做的：

数学方面：

看《数学之美》了解自然语言处理的数学基础（看了一半，还没看完）

Python方面：

1.在网上找相应的视频学习对文本进行分词处理、词频统计，用TF-IDF算法、TextRank算法、LDA建模等方法对文章进行关键词抽取

2.用朴素贝叶斯、SVM方法进行文本分类（还在看，没学完）

3.看了一下狗屁不通文章生成器的实现代码，发现实现方法其实就是随机组合语料库里的内容，现在就可以实现，可以考虑在比较简单的地方（比如新闻导语生成）进行运用，实现最简单的文本生成

下一步目标：

1.学爬虫爬取文本，首先要能自己爬取语料库并对其进行关键词抽取、完成文本分类等，其次还要能找到合适的语料库才能继续进行其他方面的工作

2.继续学习自然语言处理的相关知识（找合适的学习资源，之前找的视频感觉并不是很有用）