

本实验实现了中文分词实验的相关内容 所有的编码都统一使用**GB2312** 文件说明 其中

3.1词典的构建data_to_dic.py