1. jieba模块：
   1. with open(filepath, encode = ‘utf-8’) as f:
      1. words = jieba.lcut(f.read()) 读取文件内容并分词