1. 对语料库进行数据预处理，利用朴素贝叶斯算法或SVM完成对测试集的文本进行分类
2. 对语料库进行数据预处理，利用不同特征提取方法，如tfidf、chi2（卡方）、LDA等，利用SVM实现文本分类，并对比不同方法的准确率