

本科生实习&毕业设计（论文）工作周报						
学生姓名	乐远	联系电话	13069875982	Email	572176750@qq.com	
实习基地 (或实验室)	深交所		基地导师	徐保勋	校内导师	郭勇、张宇
报告日期	2017-8-13		实习周次	第 4 周		
报告区间	2017-8-7 到 2017-8-13					
工作重点	学习机器学习分类相关技术					
工作内容 及进度	1.学习贝叶斯分类算法的原理，并用 python 实现，并对复星系控股公司的风险状况分析，对比现成的 sklearn 库分析实验结果 2.将智能资讯系统资讯新闻模块由于算法歧义分类错误人工标注出来(一天半) 3.复查歧义分类的结果，并开会总结消除歧义的方法					
遇到问题	在分析算法进行公告分类时所产生的歧义项时碰到了很多歧义的问题，例如“鹏博士”是指股票还是指人的称呼，“三棵树”是指股票还是指有三棵树等等。怎么使用算法来消除这些歧义，更难的是怎么增强分类算法的泛化能力					
自我小结	1.学会了怎么使用贝叶斯算法进行分类，训练以及实际的应用，在使用贝叶斯分类对垃圾邮件、公告分类时都取得了不错的效果，正确率都达到了 90% 以上。 2.对歧义的消除是一个很大的难点，如果使用分类算法，那么特征提取就得做的很好；如果使用 NLP 相关方法，也许词向量、文档向量会有不错的效果。					
下周计划	继续学习机器学习相关技术，思考怎么来消除歧义					
导师意见	签字：					

## 说明：

1. 上述表格大小可以根据文字多少调整；
2. 每周要写工作周报并发送给校内导师，实习基地导师可根据需要发送；
3. 校内实习学生每 2 周要书面报呈周报给导师，当面指导并签字。