



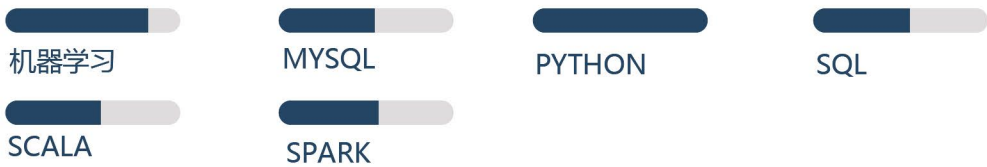
李超

求职意向：数据挖掘（机器学习）

教育背景

2015/09-2018/06	暨南大学	统计系数量经济学
2011/09-2015/06	长春大学	应用数学系信息与计算科学

IT技能



项目经历

1.小船系统大数据课题研究

项目内容：

【需求分析】结合海关业务以及现有的数据基础确立初步方案。

【统计报表系统】小船系统（针对来往港澳业务）的海关业务统计报表，警示信息，统计并将统计模块嵌入原有系统。

【轨迹数据挖掘】基于密度聚类方法规划珠江流域的各个口岸来往的主航线，以便于监测、巡查和增加航行安全（构建航道偏离报警机制）。

【警示数据筛选】警示信息与发生地理位置以及天气数据和船只属性进行关联分析，用以过滤不必要的警示。

【舱单文本数据挖掘】针对货物种类之间关联规则分析，货物地区产业模式挖掘。辅助缉私业务，提高抽样检查捕捉概率。

【项目成果】该项目被广州海关部门推广到其他地方海关。公司进一步受南宁海关邀请进行沟通交流。

角色与负责内容：主要负责建模、协调组内任务分配和进度跟进、模型code与实测、向上级汇报任务进度。

2.物流配送站点规划

项目内容：该项目属于学术性质，研究内容为给定站点距离以及GPA矩阵，配送点建设固定成本，以及各个配送点需求，以及配送给点运输可变成本，车载容量等

【模型构建】将模型数学化，合理构建目标函数（成本函数）。

【模型最优规划】通过查阅相关文章针对目标函数选择合适求解方法，最终选择遗传算法对每个配送点进行二进制编码，编写遗传算法求解最佳规划。

角色与负责内容：主要负责文献查阅，整个项目的code。

实习经历

广州实盈投资管理有限公司 量化实习生 2017/06-至今

职责：参与多因子alpha选股项目，负责大量技术、财务等因子计算，以及广泛阅读中英文研报，参与因子库构建。matrix数据API接口开发。

广州百思达数据服务有限公司 数据分析师 2016/09-2017/04

工作内容：为开发组提供算法方案，客户业务沟通与需求分析，标书撰写，帮助公司中标价值500W的项目。团队在6个月内成功拿下两个项目，广州海关信息分中心大数据建设和5184高考网站数据分析项目。公司创立与于6年2月份，在两个项目中迅速成长起来。

竞赛经历

1.广东省金融建模大赛

选题：HS300股指期货套利

【建模过程】：数据采用沪深300股指期货高频数据。通过查阅大量国内外相关文献构建反应市场行情指标体系，如能量潮、价格偏离指数等50个指标。利用PCA进行降维处理，以保证指标之间弱相关性以及剔除冗余数据。采用空间重构技术确定嵌入时间维度，SVM模型预测明日涨跌情况，以决定采取做空或做多的日交易策略，入场择时采用MACD。年化收益34%最大回撤率13%，入围优秀作品奖。

负责内容：本次参赛队员两人，本人担任组长，参与文献查阅、负责主要模型思路、并承担大部分code工作。

编程语言：python、matlab

2.研究生数学建模大赛（华为杯）

选题：致病基因挖掘

【建模过程】：数据是1w+（基因）X10w+（碱基对）基因组编码序列。剔除异常数据，如A_C碱基对，或者包含非ATCG的。特征信息是提取2-10有序碱基序列在基因中占比，并归一化数据。基于致病极有占极少数且有共同特征这一假设，利用KMeans聚类弱小的分类组确立为可疑对象。分类结果结合患病者的特征（13类）建立随机森林模型，以检验可疑基因分类能力。最终结果随机森林分类准确率超过50%以上（13类结果），小组获得优秀参赛奖。

负责内容：小组由3人组成，本人为组长，负责分配工作任务，部分思路以及code。

编程语言：R、python、matlab

感受：两次比赛都是非常宝贵经历，对后面项目帮助特别大，让我能很快将自己日常课堂课外所学有效融合到项目中。特别是在“华为杯”过程中，由于比赛时间仅仅为4天，一度接近放弃。本人考虑成员的特点合理分工明确任务，通宵2晚坚持完成。根据人员特点合理分工对一个项目的成败至关重要。

自我评价

本人学习能力和接受能力比较强、勤奋好学、有目标、热衷于人工智能行业，在工作中具有使命感，平时极大兴趣学习scala函数编程，基于spark的机器学习，以及统计学习等知识，努力成为一个成熟的数据挖掘方面人才。业余时间喜欢自己code和研究算法，写学习日记和技术博客（简书地址<http://www.jianshu.com/u/72b5343cc642>，github账户kaenleelfm@gmail.com）。

1992-03-01

湖北荆州

广州市天河区黄埔大道西601号

18122041556

lichaolfm@163.com

CTE 英语六级