

朱军
西安交通大学
西安, 中国, 710049

(+86) 187 9270 2157 (cell)
jszhujun2010@126.com
jszhujun2010.farbox.com

求职目标

数据挖掘研究员, 南京

教育背景

2014– 硕士在读, 计算机科学与技术, 西安交通大学

硕士课题: 程序bug自动修复

MOOC课程: 机器学习基石, 机器学习技法, 算法, 人工智能, 分析学的优势 (融合了机器学习, 优化问题以及数据可视化)

2010–2014 学士学位, 信息工程, 西安交通大学

班级前10名, 本硕连读班

本科毕设课题: 超图查询算法研究

项目经历

网络视频客户端

这是某个视频信号转换与转发的项目的一部分。该项目的需求是要把通过同轴电缆传输的TS流的电视信号通过MCU处理, 转换为网络流信号, 然后向计算机转发。MCU在这里还充当了web服务器的功能, 所以我实现了一个带GUI的客户端, 负责与web服务器进行通信, 并将网络流的视频信号解码播放。该客户端主要由Python实现, GUI则使用了PyQt。

Pacman游戏

该项目隶属于UC伯克利的人工智能课程。题目情景是: 有一个pacman(吃豆人)在迷宫里面, 迷宫里面有一些豆子和一些在移动的幽灵, 要求采取合适的策略, 在不同的情形下让pacman尽快吃完所有的豆子, 同时不能碰到这些幽灵。该项目由Python实现, 并应用了不少重要的搜索算法。

天池竞赛, 2015

282th/7186

这是由阿里巴巴举办的一个数据挖掘的比赛。主办方提供了2014/11/18到2017/12/17的用户购物数据, 要求预测2014/12/18的用户购买情况。该比赛分为两个阶段, 第一阶段的数据量大约为500M, 主要使用了R进行数据处理, 而第二阶段则涉及到TB数量级的数据, 则在阿里云上进行数据分析。

Kaggle竞赛

186th/2329

这是由MIT的Edx课程分析学的优势主办的关于文本挖掘的比赛。题目的设定是, 给定一些带标签的博客文本作为训练数据, 要求构建模型预测博客文章是否会火。比赛中使用了"bag of words"的模型, 并从文本中提取出了很多重要的特征, 然后尝试了不同的分类算法, 最终将一些模型集成起来。

个人技能

英语

- CET4: 601, CET6: 535

编程语言

- Python, C/C++, Java

机器学习与数据挖掘

- R, Pandas + Scikit-learn, Spark MLlib

荣誉与社会经历

思源奖学金, 2011, 2012

招生志愿者, 西安交通大学