

李蕴哲

✉ madli@ucdavis.edu · ✉ yli566@ucsc.edu · ☎ (+1) 310 795 4718 ·

教育背景

UCSC, Santa Cruz, California, USA	2018 – 至今
在读硕士研究生 贝叶斯统计, 预计 2020 年 6 月毕业	
UCD, Davis, California, USA	2016 – 2018
学士 统计	

实验室/项目经历

Data Science Initiative (Davis)	2017 年 3 月 – 2018 年 8 月
实习研究员 指导老师: Duncan Temple-Lang	
参与完善 R Package, ReadPDF	
<ul style="list-style-type: none">• 测试 Package 的合理性和提供单元测试方程。• 当 PDF 可以被转化为 XML 格式时, 提取日期信息与地点信息。• 当 PDF 为图片无法被转化成有效格式, 使用 tesseract 的 OCR 技术提取文字信息, 并用统计模型筛选有效信息。	

Kaggle 比赛

在 Santander Customer Transaction Prediction 比赛中, 使用 XGboost 和 lightGBM 算法预测用户是否会进行交易, 并获得 Leader Board 上前 15% 的成绩。

课程项目

图像处理

使用 sklearnr 里的 SVM 算法, 和 keras 的 CNN 算法判断古代西方民歌配图的人像性别。达到百分之 70 的正确率。

自然语言处理

使用 spotify 的 API 获得用户对歌曲的评论, 并使用 Naive Bayes 算法基于用户评论对歌曲进行分类判断。

IT 技能

- 编程语言: Python, R, C++, SQL
- 平台: Linux

相关课程

- 统计: 概率论, 数理统计, 时间序列, 回归分析, 方差分析, 多元统计分析, 贝叶斯统计基础, 贝叶斯统计建模
- 计算机: C++ 基础与面向对象语言, 数据结构, 脚本语言, 算法设计与分析, 数据库系统基础理论与 SQL 应用, 平行运算

其他

- GitHub: <https://github.com/Jacksoften>
- 语言: 英语 - 流利