

黄恺丹

(+86) 18664822738 | dankaihuang@outlook.com

微信: 18664822738



教育背景

悉尼大学 (QS 排名: 19) 数据科学硕士 (Master of Data Science) 2021 - 2023

主要课程: 数据科学原理、数据库管理系统、机器学习和数据挖掘、计算统计方法、
数据可视化分析、统计学习与数据挖掘、自然语言处理、社会科学数据分析

澳大利亚国立大学 (QS 排名: 34) 统计学士 (Bachelor of Statistics) 2017 - 2020

主要课程: 概率论、随机过程、回归模型、离散数学模型、经济金融模型基础、商业信息系统、
广义线性模型、图形数据分析、生存模型、关系数据库、应用数学、网页开发与设计、
贝叶斯数据分析、精算数学、大数据统计

实习经历

卡夫亨氏有限公司 (KHC) 商业分析实习生 2023.07 - 2024.01

- 使用SQL Server从数据库中提取数据并用Excel制作报表
- 协助维护销售自动化系统
- 提供技术支持, 同时定期提供周期性的分析报表
- 用PowerBI制作数据模版
- 对系统的主数据库进行维护和更新
- 同时对绩效报表和各类分析表进行测试与维护, 寻找异常数据, 并对异常数据进行检测, 找出数据异常根本原因

项目经历

联邦学习算法在图像识别中的应用项目实践 2022.07 - 2022.11

- 在学院教授的指导下利用BAGAN来恢复数据集的平衡性
- 利用联邦学习算法在基于已有的FedAvg基础上提出新模型Bi-FedAvg
- 比较和讨论FedAvg、具有条件GAN模型的FedAvg和具有BAGAN-GP模型的FedAvg在不平衡数据集的IID和non-IID环境中的训练和性能测试

对提高航空业乘客满意度的研究 2022.09 - 2022.11

- 通过R语言对航空业乘客满意度数据进行清洗和特征选择
- 运用逻辑回归、梯度提升 (Gradient Boosting)、朴素贝叶斯、随机森林和CatBoost五种方法构建模型和预测
- 使用caret包对其进行调参优化, 对训练出来的模型进行比较
- CatBoost的f1 score结果最佳, 为96.23%

悉尼娱乐与剧院 (SET) 系统开发 2021.03 - 2021.06

- 通过使用Python, PostgreSQL处理来自用户界面 (UI) 的数据库读写请求并增强数据库管理
- 创建并修改SQL架构, 以初始化SET系统数据库, 确保数据的有效组织和存储
- 开发用户认证、预订管理关键功能和通过psycpg2执行复杂的SQL查询来检索数据
- 成功显著提高了剧院预订和客户服务操作的管理效率

工作技能

- 编程: Python, R, Haskell
- 工具: SQL, Tableau, MS Office 系列办公软件, HTML, CSS, PowerBI
- 语言: 普通话, 粤语, 英语

HUANG Kaidan

黄恺丹



+86 18664822738



dankaihuang@outlook.com

Education	The University of Sydney (QS Ranking: 19)	03/2021 - 12/2023
	Master of Data Science	
	The Australian National University (QS Ranking: 34)	02/2017 - 12/2020
	Bachelor of Statistics	

Work Experience	Kraft Heinz	07/2023 - 01/2024
	Business Analyst Intern	
	<ul style="list-style-type: none">• Extract data from SQL Server database, and produce reports using Excel• Assist in maintaining a sales automation system• Provide technical support for various departments, and provide them with periodic analytics• Building data templates with Power BI• Maintain and update main database• Test and maintain KPI reports and various analytics. Identify anomalies and analyze their root causes	

Projects	Image Classification with Federated Learning	07/2022 - 11/2022
	<ul style="list-style-type: none">• Balancing data and improving the quality of image data using BAGAN.• Proposing a new model called Bi-FedAvg based on FedAvg.• Conducting experiments on FedAvg with CGAN, FedAvg with BAGANGP, and FedAvg without GAN model.	
	Research on Passenger Satisfaction in the Aviation Industry	09/2022 - 11/2022
	<ul style="list-style-type: none">• Classification project on passenger satisfaction• Using R to clean data and to conduct feature selection.• Building models with techniques such as Gradient Boosting, Naive Bayes, Random Forest, CatBoost• Using caret library to tune the hyper-parameters and make the comparison• CatBoost method achieved an f1 score of 96.23%.	
	Sydney Entertainment & Theatres (SET) System Development	03/2021 - 06/2021
	<ul style="list-style-type: none">• Using Python and PostgreSQL to handle requests for reads and writes to the database coming from the user interface and to enhance database management• Creating and modifying the SQL schema for initializing the SET system database, ensuring efficient data organization and storage.• Developing main functionalities of user authentication, booking management, and using psycopg2 for data retrieval with complex SQL queries.• Successfully improving the management of theatre bookings and customer service operations, enhancing overall efficiency	

Skills	Programming: R, Python, Haskell
	Technology: SQL, Tableau, MS Office, HTML, CSS, Power BI
	Languages: Mandarin, Cantonese, English