

# 吴嘉澍

男 1997.06.13 邮箱 [jiashu@student.unimelb.edu.au](mailto:jiashu@student.unimelb.edu.au) 微信 [jiashu613](https://www.wechat.com/p/jiashu613) [linkedin.com/in/jiashu-wu](https://www.linkedin.com/in/jiashu-wu)  
籍贯 北京 海淀 电话 17801323125 Skype [live:davidjiashuwu](https://www.skype.com/p/live:davidjiashuwu) [github.com/jiashuwu](https://github.com/jiashuwu)

## 教育背景

**墨尔本大学** 信息技术硕士 人工智能 2019.02 – 2020 年 12 月毕业  
均分 88.1 专业前 2% 全科 First Class Honours 入选 2019 Dean's Honours List  
**悉尼大学** 理科学士 计算机科学 金融数学与统计学 双专业 2016.02 – 2018.12  
均分 86.5 专业前 2% 计算机科学均分 85 金融数学与统计学均分 88 入选 2017 & 2018 Dean's Honours List  
**北京理工大学** 软件工程 2015.08 – 2016.01  
软件工程专业 2016 年转学至悉尼大学就读

## 实习经历

**北京理工大学计算机学院** 研究实习生 2019.11 – 2020.05

课题: Simultaneous Semantic Alignment Network for Heterogeneous Domain Adaptation

- 探究挑战性大的 HDA 问题, 难点在于特征表达与模态的异构性。通过不断试验分析, 成功克服 HDA 模型标签准确率低等问题。
- 负责探索不同的网络结构, 试验不同的迁移方法。通过主动思考, 最终有效地使用知识蒸馏、中心点匹配、标签精细化等方法, 实现对于模型迁移能力的优化增强。成功地使模型在多个 HDA 数据集上比现有 HDA 方法均有约 1~6% 的预测准确率提升。
- 负责使用 PyTorch 深度学习框架实现模型, 并利用 Git 维护管理不同版本的模型代码。代码可读性高, 注释清晰, 整洁易懂。
- 在有限时间内高质量完成 9 千字英文论文。论文及模型创新性获审稿人一致认可, 并已被 CCF-A 类会议 ACM MM'2020 接收。

## 项目经历

**墨尔本大学 Profware 学者搜索平台** 2019.07 – 2019.11

课题: Learning to rank with small set of ground truth data

- 负责开发学者搜索平台。不惧难点, 在规定时间内高效试验矩阵分解及知识库两种解决方案, 成功解决排序数据不足的难点。
- 运用 Python (NLTK, spaCy) 对 40 万条论文数据及词条数约为 2 千万的 Wiki 和 MAG 词库进行数据预处理, 提取与整合。
- 缜密分析并对比多种协同过滤方法的优劣, 包括 LSA, NMF 及深度推荐系统算法 NFM。高效处理、挖掘大量数据, 利用数据量约 20 亿的 MAG Knowledge Base 辅助实现搜索结果的推荐及排序, 成功解决了搜索词未在论文中出现等难点。
- 高质量撰写 12000 字近 50 页的报告, 详细展示了数据获取, 数据处理, 模型试验结果, 最终模型方案等内容, 并做汇报。

**Kaggle 推特文本作者归属分类** 2019.07 – 2019.09

- 利用 Python 对 30 万条推特进行数据清洗和预处理。开展特征工程, 并进行特征选择。利用 SMOTE 算法解决样本不均衡的问题。
- 负责使用 SVM, 随机森林, TextCNN, TextRNN 等算法对数据进行分类, 成功将准确率提升至 30%, 排名 20 (约 200 支队伍)。

## 专业技能

- 熟练使用 Python (NumPy, NLTK, Matplotlib, Sklearn, PyTorch, Pandas, PyMySQL 等), Java (JDBC), SQL, R(假设检测, 统计分析, 作图, 时间序列分析等), 可使用 Haskell, C。自学 Hadoop, Spark 及 Hive 基础以及 Power BI 可视化工具。
- 熟练使用 Windows 操作系统, Microsoft Excel, PPT, Word, Visio, 等软件, 熟悉 Linux 操作系统。可使用 Git 进行版本管理。
- 熟悉数据清洗, 数据处理, 数据建模, A/B Testing 等数据处理分析流程。熟悉聚类, 神经网络等多种机器学习、深度学习算法。
- 主要课程: 机器学习, 数据分析, 统计学习, 数据库, 自然语言处理, 统计建模, 统计测试等。
- 雅思学术类 (总分 7.0, 阅读 8.0, 听力 8.0, 口语 6.0, 写作 6.0)。大学英语四级 (总分 665, 阅读满分)。在澳大利亚五年学习生活经历, 熟练适应英语交流环境, 能够熟练使用英语进行汇报及写作, 具有良好的沟通交流表达能力。

## 其他经历

- 具有极强的学习能力。从零自学安卓开发, 业余时间编写的倒数日记事本 App 已上架华为、阿里等应用市场, 已有几千次下载量。
- 善于将技术应用于实践。在悉尼大学 MHTI 实验室负责编写收集用户绘制图形的笔迹坐标及触点压力数据的 App, 数据采用 SQLite3 进行高效存储。应用可靠易用, 帮助实验室顺利完成数据采集并用于帕金森诊断的研究, 后续研究进展顺利。
- 担任北京理工大学学生会体育部部长, 责任心强, 参与筹备上千人参与的校运会及长跑。协调沟通场地等多部门, 活动顺利举办。
- 担任悉尼大学网页设计, Java 程序设计及数据结构课助教, 为上百名同学答疑, 连续两学期获得学院优秀助教荣誉。