



# 和瑞 Rui He

机器学习方向博士在读 · 伯明翰大学 & 南方科技大学

深圳, 中国

☎ (+86) 185-6577-0048 | ✉ her2018@mail.sustech.edu.cn | 📅 1996 | 🌐 www.superui.cc | 📷 SupeRuier

## 个人陈述

我是伯明翰大学 (UoB) 和南方科技大学 (SUSTech) 联合培养的一名富有探索精神和自驱力的博士研究生。在计算机科学和物理学方面, 我有着扎实的学术基础。我的研究工作主要集中在 **机器学习** 方面, 特别是在 **主动学习** 和 **多领域学习** 上。我独立高效学习的能力, 清晰的逻辑思维, 以问题为导向的思维方式带来了高质量的研究成果。在 AI 领域, 我曾多次担任本科生以及研究生的课程助教, 在授课中我展现出了优秀的知识组织能力以及授课热情。当下, AI 的快速发展及其在实践中的广泛应用给我带来了强烈的震撼, 在此我寻求一个充分利用我 AI 方向的知识技术储备解决现实中的重大问题的具有挑战性的角色, 以期实现科研与实践的完美结合。

## 教育经历

### 伯明翰大学

伯明翰, 英国 & 深圳, 中国

**计算机科学与工程博士在读** (南方科技大学联合培养项目)

2018 年 9 月 - 2023 年 12 月 (预计)

- ◇ 导师: 唐珂教授 (南方科技大学), 何山教授 (伯明翰大学)
- ◇ 研究方向: 主动学习, 多领域学习, 域迁移等
- ◇ 毕业论文题目: Advancing Multi-Domain Active Learning: Strategies and Techniques Using Neural Networks
- ◇ 关键成长: 在读博期间养成了以问题导向思维的思维方式。同时在科研训练中锻炼了逻辑思维能力, 以及信息搜集能力。

### 南方科技大学

深圳, 中国

**物理学学士** GPA:3.65/4

2014 年 9 月 - 2018 年 7 月

- ◇ 导师: 徐虎教授, 卢海舟教授
- ◇ 课程: 量子力学, 统计物理, 概率论与数理统计, 数据结构与分析等
- ◇ 荣誉: 优秀新生奖学金, 政府初创奖学金, 树仁书院杰出毕业生
- ◇ 关键成长: 锻炼了独立快速学习并掌握一个新方向的能力。

### 天普大学

费城, 美国

**学期交换项目** GPA:3.93/4

2016 年 8 月 - 2016 年 12 月

- ◇ 课程: 科学计算方法, 原子核与粒子物理, 问题导向的 Python 编程等
- ◇ 关键成长: 克服对陌生环境的不适应感, 编程能力大幅提升, 英语交流能力极大提高。

## 学术论文

2022 **Multi-Domain Active Learning: Literature Review and Comparative Study**, First author

期刊论文 - IEEE Transactions on Emerging Topics in Computational Intelligence (TETCI)

- ◇ 在多领域 (不同来源不同分布) 的数据上构建模型是现实世界中十分重要的问题, 但其仍面临着标注成本昂贵的挑战。
- ◇ 我们提出了一个将主动学习和多领域学习结合起来的多领域主动学习 (MDAL) 解决方案以降低标注成本。
- ◇ 我们提供了 MDAL 的正式定义以及第一篇文献综述, 并在我们提出的流程下进行了全面彻底的对比实验。

2023 **Multi-Domain Learning From Insufficient Annotations**, First author

会议论文 - 26th European Conference on Artificial Intelligence (ECAI)

- ◇ 在现实世界中存在从不充足的标注数据中进行多领域学习的需求, 但是此场景尚未得到充分的研究。
- ◇ 我们提出了一种基于对比学习的多领域学习方法, 同时利用标注数据和无标注数据优化表征, 提升少标记下模型表现。

2023 **Perturbation-Based Two-Stage Multi-Domain Active Learning**, First author

会议论文 - 32nd ACM International Conference on Information and Knowledge Management (CIKM)

- ◇ 传统的主动学习策略由于忽略了跨领域信息的评估, 因此在多领域主动学习下通常会得到次优选取。
- ◇ 我们提出了一种基于扰动的两阶段 MDAL 方法, 直接评估了跨领域信息以进一步选取高信息量的样本进行标注。

2023 **Large Language Models can be Guided to Evade AI-Generated Text Detection**,

- ◇ 我们提出了一种基于上下文学习的文本攻击方法, 以指导 LLMs 生成具有更少模型特征 (水印/机器回复) 的文本。

- ◇ 我们提出了一种基于双层优化的推荐数据蒸馏方法，基于大量用户行为数据生成少量高质训练样本，以快速训练模型。

## 个人技能

- 软技能** 以问题导向的思考方式，以结果为导向的工作习惯，快速学习及知识构建能力，较强信息搜集能力，较强教学能力
- 编程技能** Python (PyTorch, Numpy, Matplotlib, Scikit-learn, etc.), Java,  $\text{\LaTeX}$
- AI 大模型** Hugging face, ChatGPT API, Stable Diffusion
- 开发** Hexo (static website generator), Git, Nginx + Tencent Cloud (web page deployment)
- 语言能力** 中文 (母语), 英文 (高级商务沟通)

## 工作经历

### 南方科技大学

深圳, 中国

博士生助教 - 智能数据分析 CSE5002 (每年春季学期)

2020 年 1 月 - 2023 年 7 月

- ◇ **负责研究生实验课授课**, 每学期 32 学时, 内容覆盖机器学习相关知识点, 包含**模型训练**, **数据处理**, **特征工程**, **非凸优化**等。
- ◇ 实验课材料准备 (讲义及演示代码)。此外负责课程作业设计与学期课程编程项目设计。
- ◇ **关键性个人成长**: 在本门课多次授课过程中, 深度训练了授课和沟通能力, 也提升了知识串联理解能力。

博士生助教 - 人工智能 CS303 (每年秋季学期)

2019 年 9 月 - 2022 年 1 月

- ◇ **负责本科国际生实验课远程授课**, 主要包含**搜索**, **优化**, **机器学习**等基础知识。
- ◇ 实验课材料准备 (编程演示代码、Python 科学计算基础知识培训), 及实验课答疑。
- ◇ **关键性个人成长**: 在与国际生交流过程中, 培养了远程授课能力以及远程技术支持技巧。

### 伯明翰大学

伯明翰, 英国

博士生助教 - Data Structures and Algorithms

2019 年 2 月 - 2019 年 6 月

- ◇ 面向伯明翰大学本科生, 负责实验课编程项目的讲解, 每周两课时。此外课后负责编程作业的讲解和批改。
- ◇ **关键性个人成长**: 培养了与授课教师以及学生的沟通技巧。

博士生助教 - Software Workshop (Java)

2018 年 10 月 - 2018 年 12 月

- ◇ 面向伯明翰大学研究生, **负责 Java 实验课授课** (知识点教授以及案例分析), 每周两课时。此外负责编程作业批改。
- ◇ **关键性个人成长**: 极大地培养了英文授课能力, 以及提升了公开演讲能力。

## 项目经历

### Awesome Active Learning

Github 开源仓库

创立者 & 维护者

2020 年 2 月 - PRESENT

- ◇ 本项目是一个结构严谨的, 包含全方面内容的主动学习知识库。旨在使用户于一处即可检索到关于主动学习的一切重要资料。
- ◇ 项目包含 AL 研究方向下的分类学逻辑, 最新相关论文, 成熟的代码, 以及其他有价值的资料。
- ◇ 至 2023 年八月, 约有 500 stars。

### 多领域主动学习: 一个框架

Github 开源仓库

创立者 & 开发者

2022 年 1 月 - 2022 年 12 月

- ◇ 一个多领域主动学习的通用框架, 面向 MDAL 问题可以直接结合不同的多领域模型以及主动学习策略作为解决方案。

## 其他事项

### 心肺复苏以及自动体外除颤仪资格证

中国红十字会

- ◇ 依照《中国红十字会法》培训及考核合格, 满足实施心肺复苏以及使用自动体外除颤仪资质。