

# 曹诗磊

手机：(+86) 13413674036 · 邮箱：j799017232@gmail.com

2024 届 · 主页：shileicao.github.io



## 教育背景

中山大学 人工智能 本科

2020.09 - 2024.06

- 校园荣誉：美国大学生数学建模竞赛一等奖 (2023)、中山大学三等优秀学生奖学金 (2021、2022)、亚太地区大学生数学建模竞赛三等奖 (2022)
- 主修课程：复变函数 (99)、线性代数 (98)、数学分析 (97)、Matlab 语言与应用 (94)、认知科学与类脑实验 (92)、机器学习 (92)、高级程序设计与方法 (92)、人工智能原理 (91)、操作系统原理 (91)、模式识别与计算机视觉 (88.5)、自然语言处理 (87)、数据结构与算法 (87)

## 科研经历

批判性论文阅读训练工具 中山大学人工智能学院人机交互实验室

2022.09-2023.04

岗位: 科研助理 导师: 彭振辉助理教授

- 以共同一作身份完成论文, 已被 **UIST 会议 (CCF-A)** 接收
- 项目描述: 研究开发一个在阅读论文时**训练批判性思考能力**的工具, 其通过利用**总结**来帮助读者掌握论文的**主要思想**, 并利用基于模板生成的**批判性问题**来帮助他们学习如何进行**批判性思考**
- 主要工作:
  - **形成性研究**: 进行相关论文调研提炼, 且线下采访 50+ 调研对象, 收集及分析科研新手在批判性阅读过程中会遇到的**困难点**; 基于论文调研和访谈结果, 得出工具的**设计需求**
  - **文本总结模型**: 通过部署 **BRIO** 模型来生成论文内容的**参考总结**, 然后利用 **Bert** 模型对原文和文本摘要的句子进行**语义相似性匹配**, 来提示用户的总结在原文中可能遗漏或者冗余的内容
  - **批判性问题生成模型**: 同时使用句子分类模型对句子分类, 再通过 **YAKE** 模型进行关键词提取后, 根据句子的类别填入经过我们模板化的批判性问题模板后生成参考的批判性问题
  - **后端实现**: 后端通过利用 **Flask** 框架与前端进行交互, 根据请求调用相关模型并返回结果, 并储存相关记录, 从而实现了工具的后端
  - **用户实验及数据分析**: 将 24 个参与者分成两组后进行**对比实验**, 对实验数据进行**曼惠特尼检验**和**二维混合方差检验**后得出结论: 与另外基线工具对比, 这个工具更能促进用户的**批判性思考**

## 论文专利

1. **CriTrainer: An Adaptive Training Tool for Critical Paper Reading**  
Kangyu Yuan\*, Hehai Lin\*, **Shilei Cao\***, Zhenhui Peng, Qingyu Guo, and Xiaojuan Ma (\* 同等贡献)  
ACM Symposium on User Interface Software and Technology (UIST), 2023 (**CCF-A**)
2. 单云霄, 李国豪, 张俊东, **曹诗磊**, 杜心龙, 何子杰, 邱炜. 基于椭圆鲁棒性控制的多机器人协同追捕方法及系统. 2023101368382. (2023, 专利正在申请, 目前已公开)
3. 单云霄, 白昊天, 李国豪, **曹诗磊**, 杜心龙, 何子杰, 邱炜. 一种基于泰森多边形的多机器人追逃避障的方法及系统. 2023101532796. (2023, 专利正在申请, 目前已公开)

## 专业技能

- **深度学习**: 熟练掌握深度学习的基本原理, 包括 CNN、RNN、LSTM、GAN、Transformer、Bert、GPT 等模型。具备使用深度学习框架 PyTorch 进行模型搭建、训练和测试的实践经验。熟悉常见的深度学习模型优化策略, 如批量归一化、dropout 等
- **数据分析**: 熟练掌握 Python 数据分析库, 如 NumPy、Pandas、Matplotlib 等, 具备数据清洗、处理、可视化等技能。具备基本的统计学知识, 能够运用 SPSS 进行数据分析
- **机器学习**: 了解常见的监督学习和非监督学习算法, 如线性回归、支持向量机、决策树、聚类等。能够使用机器学习库 (如 scikit-learn) 完成数据预处理、特征工程、模型训练和评估等任务
- **后端开发**: 熟悉并掌握 C++、Python 等开发语言, 理解并掌握基本的数据结构与算法, 有良好的软件开发能力
- **语言能力**: CET-6 (553), 口语表达较流畅标准; 母语普通话、粤语