



🎓 教育背景

中国科学技术大学 (USTC) 本科 人工智能与数据科学学院 数据科学与大数据技术 2022.9-2026.6

绩点: 3.91/4.3(90.74/100) 专业排名: 4/51

- 专业课程: 线性代数B1 (A+)、离散数学 (A+)、运筹学 (A)、数学分析B1 (A)、几何学基础 (A)、复变函数B (A)、概率论与数理统计 (A-)、计算方法 (A-)、人工智能原理与技术 (A+)、数据分析编程导论 (A+)、Python科学计算编程基础 (A)、数据结构 (A)、计算系统导论A (A)、信号与系统 (A)
- 科研特色课程: 科学与社会研讨课 (A+)
- 四级 598 六级 489 (口语优秀) 具备独立阅读与撰写英文论文的能力

🏆 获奖情况

- 奖学金类:
 - 优秀新生奖学金银奖 (前15%) USTC, 2022.11
 - 基础学科拔尖计划英才班 (严济慈物理英才班) USTC, 2022.11
 - 地奥奖学金 USTC, 2023.11
 - 华为奖学金 USTC, 2024.11
 - 国家励志奖学金 USTC, 2024.11
- 竞赛类:
 - 第十五届全国大学生数学竞赛二等奖 CMS, 2024.1

🔧 工作技能

- 编程语言: Python、C、C++、R
- 脚本语言: \LaTeX 、Bash、HTML、CSS、JavaScript
- 工具与技术: Git、Docker、SQL、Jupyter、Torch、Django

🏢 科研经历

【大语言模型安全研究】科研实习 2024.12至今

- 项目概述: 在张岸教授 (时任NUS Next++实验室Research Fellow, 现中国科学技术大学特任教授) 指导下, 参与大语言模型 (LLM) 安全与鲁棒性研究, 聚焦模型拒绝机制与防护策略优化。
- 负责工作: 设计并实现基于零空间约束的拒绝向量方法, 分析模型在恶意提示下的行为, 优化LLM的安全性与性能; 与NUS Next++实验室合作, 参与方法设计与实验, 撰写部分研究论文。
- 项目成果: 以共同一作身份投稿NIPS 2025, 论文《AlphaSteer: Learning Refusal Steering with Principled Null-Space Constraint》(arXiv:2506.08390), 提出创新方法提升LLM安全性, 填补相关研究空白。

🏠 项目经历

【iGEM国际基因工程机器大赛】代码组组长 2024.10

- 项目概述: 领导USTC-Software团队参与iGEM国际基因工程机器大赛, 开发前后端平台, 提升数据可视化与分析功能。
- 负责工作: 领导代码组, 负责前端与后端开发, 确保系统无缝整合与高性能; 部署机器学习模型以增强平台功能; 在巴黎iGEM国际基因工程机器大赛中展示平台功能, 与国际团队交流并收集反馈。
- 项目成果: 成功展示平台功能, 获得国际团队的积极反馈, 推动平台持续优化。

【香港科技大学物理系交流项目】 2023.8

- **项目概述:** 参与为期一周的交流项目, 深入了解机器学习与物理学的交叉研究, 参加 “When Physics Meets Machine Learning” 系列讲座。
- **负责工作:** 在香港科技大学专家指导下, 探索机器学习与物理学的前沿研究与进展。
- **项目成果:** 深入了解跨学科研究, 拓宽学术视野。

【本科课程项目】

- **非线性动态系统模拟** 2023.11
 - 使用Python模拟非线性动态系统, 分析其行为并对结果进行可视化, 设计UI简化控制变量分析过程。
- **五子棋模型** 2024.6
 - 与同学合作使用深度强化学习 (DRL) 设计五子棋模型, 击败助教提供的Baseline。
- **英语俱乐部交流平台** 2024.7
 - 领导两人团队, 使用Django框架开发英语俱乐部交流平台, 实现多种功能, 并提供用户友好的交互界面。
- **校园文件资源检索系统** 2024.12
 - 领导两人团队, 开发校园文件资源检索系统, 集成检索增强生成 (RAG) 技术, 尝试优化本地部署的聊天机器人模型。

课外活动

- **中国科学技术大学学生英语俱乐部, 社长** 2023.9-2024.9
 - 规划并协调社团活动, 总结工作并评估成果。
 - 组织社团历史上最大规模的户外寻宝活动——USTC Mystery Hunt。
 - 管理社团官方账号, 设计宣传材料, 撰写宣传文章。
- **中国科学技术大学物理系篮球队, 队员** 2022.9至今
 - 积极参加校内篮球比赛, 包括新生篮球锦标赛、科星杯和四国大战等。
- **健身、篮球、跆拳道俱乐部, 成员** 2022.9至今
 - 参与各类健身与运动活动, 倡导健康积极的生活方式。
- **中国科学技术大学主持俱乐部, 成员** 2022.9至今
 - 主持多项校级活动, 包括 “美丽邂逅”、“舞星”、“春见”、“溯光” 等。
 - 参加演讲与朗诵比赛并获奖, 具有优秀的公众演讲能力。

教学助理经历

- **离散数学 (本科课程)** 2024秋季
- **线性代数B1 (本科课程)** 2025春季
- **人工智能与机器学习基础 (本科课程)** 2025秋季

论文

* 同等贡献, † 通讯作者

- **预印版**
 - Leheng Sheng*, **Changshuo Shen***, Weixiang Zhao, Junfeng Fang, Xiaohao Liu, Zhenkai Liang, Xiang Wang, An Zhang†, Tat-Seng Chua, “AlphaSteer: Learning Refusal Steering with Principled Null-Space Constraint,” arXiv:2506.07022
 - Leheng Sheng, An Zhang†, Zijian Wu, Weixiang Zhao, **Changshuo Shen**, Yi Zhang, Xiang Wang, Tat-Seng Chua, “Title Not Provided,” arXiv:2506.08390
- **已发表**
 - (暂无)