



吴承儒

联系电话: 15881038217

籍贯: 四川

邮箱: chengru_wu@buaa.edu.cn

政治面貌: 共青团员



教育背景

2023.09-至今 北京航空航天大学 沈元学院 计算机拔尖计划

外语水平: CET-6 571 分 CET-4 635 分

核心课程: 高等代数(1) (99)、数学分析(1) (94)、高等代数(2) (92)、数学分析(2) (96)、离散数学(1) (95)、工科概率统计 (99)、高级语言程序设计 (99)、数据结构与程序设计 (93)、计算机组成 (95)、操作系统 (94)

专业能力: 熟练掌握 C、C++、Java 语言; 牢固掌握计算机学科和数理基础知识; 有一定的科学研究能力, 能够在专业竞赛及科研中进行实践应用。

科研竞赛

2026.01-2026.05 星海图 (北京) 人工智能科技股份有限公司——前沿算法组实习

深度参与 G0.5 VLA (Vision-Language-Action) 模型预训练, 全栈式负责核心训练数据的标注、质检与清洗验证。熟悉具身智能落地闭环, 具备高质量数据构建、模型训练调优到机器人真机部署的端到端实战经验。

2024.07-2026.01 复杂关键软件环境全国重点实验室——张莉老师课题组实习

深度参与课题组代码生成方向研究, 累计发表/录用论文 4 篇。

- 成果: 在 Internerware 2025 以学生第二作者发表论文:

AdaptiveLLM: A Framework for Selecting Optimal Cost-Efficient LLM for Code-Generation Based on CoT Length

简介: 针对代码生成任务中性能与成本的平衡难题, 我们提出以思维链 (CoT) 长度量化任务难度, 将 CodeBERT 作为感知层、XGBoost 作为决策层, 构建自适应模型选择框架, 保证性能的同时降低模型成本。

个人贡献: 负责核心训练数据的构建与清洗, 设计并执行了多模型对比实验, 验证了框架在降低推理成本同时保持高性能的有效性, 参与论文写作工作。

- 成果: 在 TSE 2025 以学生第三作者发表论文:

On the Applicability of Code Language Models to Scientific Computing Programs

简介: 为了填补了 LLM 在科学计算语言 (MATLAB/R) 研究的空白, 我们构建了涵盖算法、数学计算及 API 调用的多语言数据集 SC-CODE, 并且通过多种方法研究如何提高大模型在科学计算编程语言的表现。

个人贡献: 承担了 Baseline 模型的大规模推理实验, 通过微调 (Fine-tuning) 与少样本学习 (Few-shot) 实验分析, 揭示了通用大模型在特定科学领域的局限性。

- 成果: 在 TOSEM 2026 以学生第三作者发表论文:

Towards Realistic Project-Level Code Generation via Multi-Agent Collaboration and Semantic Architecture Modeling

简介: 针对现有模型难以生成复杂项目结构的问题, 提出“架构设计-骨架生成-代码填充”的三阶段多智能体协作框架, 并引入语义架构树与记忆管理机制。

个人贡献: 收集数据并构建了项目级代码评估基准 CODEPROJECTEVAL, 复现了现有 SOTA 模型和方法并完成对比评测, 为论文提供了关键的实验数据支撑, 参与论文数据集构建和实验结果分析的写作工作。

- 成果: 在 ESEM 2026 以学生第二作者发表论文:

Can LLM Coding Assistants Support Emerging Programming Languages? An Empirical Study on Cangjie

简介: 构建含 248 道任务的仓颉语言基准 CangjieBench, 评测六个大模型在直接提示、语法约束提示、文档检索、代码检索和智能体五种方法下的表现, 比较生成准确率、编译率与 token 成本。

个人贡献: 参与了评估基准验证与实验研究, 重点关注不同适配方法的实验比较和结果分析, 为论文绘图并参与实验部分写作。

2024.09-2025.09 第十八届市级大创: 低成本可控换装行人虚拟数据集生成技术研究

立项人

- 成果: 成功结项并在 ICME 2025 以第二作者发表论文:

CCUP: A Controllable Synthetic Data Generation Pipeline for Pretraining Cloth-Changing Person Re-Identification Models

简介: 针对行人重识别 (Re-ID) 数据匮乏与隐私问题, 我们设计了一套包含骨骼生成、监控模拟、行人检测的低成本虚拟数据生成流水线, 并通过实验证明数据集对重识别模型的提升显著。

个人贡献: 作为立项人统筹项目进度。主导生成了目前规模最大的换装行人虚拟数据集, 实验证明该数据显著提升了 Re-ID 模型在现实场景中的泛化能力。

2024.01、2025.01 美国大学生数学建模竞赛

成员 (不区分)

- 成果: 获 2025 年美国大学生数学建模竞赛 M 奖、2024 年美国大学生数学建模竞赛 H 奖

其他获奖

荣誉类: 2025 国家奖学金

2024 国家奖学金

北航学习优秀奖学金特等奖 (2 次)

北航校级三好学生

北航学科竞赛奖学金特等奖

北航优秀运动员奖学金二等奖 (2 次)

自我评价

专业知识扎实: 掌握本专业基本理论知识, 并能与实践结合。

良好实践能力: 多次参与专业竞赛及科研, 实践经历丰富。