

马周原

(+86) 13069127150 · mazy2121@mails.jlu.edu.cn · @github-page



教育背景

吉林大学 - 计算机科学与技术学院 - 唐敖庆班理科试验班 (拔尖计划) 2021.9 - 2025.6

- **GPA:** 3.90 / 4.00, 平均分数: 92.86 / 100, 专业排名: 5/18 (已确定有保研资格)
- **核心专业课成绩:** 机器学习与 Python(96); 人工智能原理 (90); 数据结构 (95);
- **奖学金荣誉:** 新生入学奖学金、三等奖学金 * 2 (大一、大二学年, 专业排名前 5).

相关能力

- **英语水平:** 英语四级 (CET-4) 607 分; 英语六级 (CET-6) 576 分
- **代码能力:** 掌握 C, C++, python 等语言; 熟悉 pytorch 架构; 熟练使用 Git 等工具; 熟悉 Linux 环境配置.
- **竞赛成绩:** CSP 认证 290 分 (累计排名 5.22%); 数学建模国赛、美赛、蓝桥杯编程能力大赛奖项若干.
- **开发经验:** 熟悉 QT 开发.

项目经历

项目 **Lead** | 基于检索增强的生成模型研究 2024.1 - NOW

- 项目负责人
- 研究方向为检索增强生成 (Retrieval-Argumented-Generation) 技术, 主要研究 Post-Retrieval 策略中的 Re-ranker 训练.
- 针对 Re-ranker 在 Hard-Negative 上表现不佳的难题, 设计了一种基于 LLM 的对抗训练 Hard-Negative-Mining 方案, 可使 Re-ranker 更加关注语义相关性的同时, 逐步学习分辨越来越困难的负样本.
- 目前已在 PopQA 数据集上取得超过 SOTA 的表现, 计划进一步在更多数据集上验证.

国家级大创项目 | 多模态语音诊断 2023.5 - 2023.11

- 国家级大创项目负责人, 主导项目的选题、pipeline 的构建、以及实验.
- 研究方向为多模态语音抑郁诊断, 主要研究文本 + 音频的多模态抑郁症分类模型.
- 针对现有抑郁诊疗模型中存在的模态融合、弱势模态等问题, 进行了调研、探索和实验.
- 相关的探索、成果、代码开源在 @github.

科研讨论班 | 医疗图像分割 2023.1 - 2023.5

- 在老师的指导下, 系统学习了 DL 经典方法和 CV 领域经典模型, 包括但不限于 GAN、Transformer、CLIP 等模型及其变种.
- 在学习过程中, 阅读了相关多篇经典 paper, 并在感兴趣的任务上进行了探索和实验, 相关的学习记录和代码开源在 @github.
- 在讨论班中, 积极参与学术讨论和论文分享, 培养了全面的文献阅读和汇报能力.

个人总结

- 本人自我驱动能力强, 乐观向上, 有强烈的集体荣誉感和团队精神, 同时对探索新领域和建立新知识体系充满热情.
- 本人热爱科研, 想攻读博士学位. 本科期间在 CV、MM、LLM 等领域的探索学习激发了我对科研的兴趣, 也让我认识到科研需要深入思考、坚持不懈、厚积薄发. 希望在后续的科研学习能中不断探索自己和知识的边界, 做有意思、有意义、有深度的科学研究.
- 现阶段我对多模态学习和 LLM 微调/应用十分感兴趣, 希望未来可以在类似领域继续学习和深造, 同时也有探索更多前沿创新领域的热情.