

# 余家康

求职意向：研究型实习生/算法实习生



男 | 23岁 | 15527708770 | JiakangYu\_JHUN | jiakangyu@stu.jhun.edu.cn

## 教育经历

江汉大学 硕士 电子信息-人工智能 2024-2027

个人学术主页: <https://kaka-yjk.github.io>

近期研究方向: 多模态虚假信息检测、多模态情感分析、手机GUI Agent

GPA和奖学金: 获一等学业奖学金 (Top 1), 二等学业奖学金 (20%), GPA 3.45

一战考研成绩: 初试373分 (数学119), 本专业初试成绩和总成绩第一 (1 / 71)

实践实训经历: 作为研究生新生代表发言, 加入武汉市开源创新中心 (人工智能方向)、CSDN开源与智能技术创新实验室

江汉大学 本科 电子信息工程 2020-2024

GPA和奖学金: GPA 3.25 (30%), 获单项奖学金

创新创业经历: 担任大学生创新创业项目负责人, 带领组员完成高质量的商业计划书, 受邀参与武汉经开区政策宣讲项目路演活动, 并完成各个模块基础功能实现

社团干部经历: 担任江汉大学机器人爱好者协会社长, 负责管理实验室设备, 组织十余次社团培训

## 专业技能

- 具备良好的文献阅读能力、论文复现能力、idea创新能力、论文写作能力。一篇一作B会在审, 一篇一作C会已中。
- 熟悉Python、Java编程及Pytorch、DeepSpeed框架, 熟悉深度学习、机器学习、强化学习理论, 有一定的调参经验。
- 有VLM智能体RFT、LLM的预训练、指令调优、DPO、蒸馏、Prompt设计的经历。有大数据开发、后端开发的项目经历。

## 科研经历

通过强化选择与对齐的多模态大模型理据以增强情感分析 2025.06-至今

我们的论文已投稿ICASSP 2026 (CCF-B)。(第一作者)

该项目已申请国家发明专利并被受理。

任务背景: 多模态情感分析是从一个包括MP4格式音视频和对应转录文本的样本中预测出[-3,3]的回归值。值越大表示越积极, 值越小表示越消极。

关键词: 多模态大语言模型, 强化学习, 缓解幻觉

摘要: 利用MLLM生成文本解释, 为增强多模态情感分析提供了一条有前景的途径。然而, MLLM所生成解释的实用性受到两大关键挑战的限制: 一是模型在选择和整合能力上的不足, 二是模型因“幻觉”问题导致其生成内容的不可靠性。为应对这一挑战, 我们引入了一个名为SAFE (用于多模态情感分析的选择性、对齐性与融合性解释) 的新颖框架, 旨在系统性地选择、对齐并融合这些由MLLM生成的解释。SAFE 采用了一个三阶段流水线: 一个强化学习模块负责挑选最关键的解释; 一个对齐模块将这些解释与源模态特征进行对齐, 以过滤幻觉内容; 一个融合模块则将经过筛选和对齐的解释与源模态特征相整合, 以指导最终的预测。在两个公开的多模态情感分析基准数据集上进行的大量实验证明, SAFE 在 CMU-MOSI 上取得了SOTA的结果, 并在 CH-SIMS 上表现出具有竞争力的性能。我们的代码发布在<https://github.com/kaka-yjk/safe>。

我们的论文已被PRICAI 2025 (CCF-C, 中稿率30%) 接收, Oral Presentation。(第一作者)

任务背景: 如上。

关键词: 多模态融合, 跨模态对齐, 对比学习

摘要: 多模态情感分析需要能够同时捕捉跨模态一致性和模态内差异的鲁棒表示。现有的融合方法通常无法适应多样化的情感线索并忽略了模态间的相关性, 而对比学习方法则对样本对的分布和损失函数的设计考虑不足。我们提出了一种用于多模态情感分析的自适应多尺度卷积融合与对比学习网络 (AMCCL), 该网络利用一个自适应多尺度卷积 (AMC) 模块来动态地融合多模态信息。AMC模块通过带有自适应加权的多尺度卷积和squeeze-and-excitation模块来动态融合特征, 以增强显著的特征通道。我们提出的细粒度对比学习利用了情感的极性和强度, 并通过定制化的损失函数来强化正样本对, 并平衡模态间与模态内的关系。在 MOSI 和 MOSEI 基准数据集上的实验表明, 与当前最先进的方法相比, AMCCL 取得了具有竞争力的性能。

## 实习经历

湖北星纪魅族集团有限公司 大模型算法实习生

2025.09-2025.10

实习内容: 主要聚焦于手机端 GUI Agent 的训练优化工作, 核心负责强化微调 (RFT) 及配套数据预处理流程。基于 AgentCPM-GUI 8B 多模态大模型和 11 万条训练样本, 在模型训练中采用 GRPO 算法, 通过对模型生成的多组交互响应进行分组比较与方差缩减优势计算, 优化模型在跨应用手机 GUI 场景下每一步的逻辑推理能力。同时, 设计了  $\{-1, 0, 1\}$  的离散奖励机制: 首先校验输出格式的正确性 (格式错误赋值 -1), 随后判断动作类型与位置的准确性 (如 POINT 坐标是否落在真实边界框内, 正确赋值 1、错误赋值 0), 以此精确引导模型的单步动作学习。在数据预处理环节, 对公司数据中的冗余轨迹进行去重与质量筛选, 确保训练数据的有效性。上述工作直接提升了 GUI Agent 在手机端长周期交互、跨界面操作等复杂任务中的表现, 同时也加深了对移动端 GUI Agent从数据处理到强化训练的完整流程理解。

## 项目经历

快餐大数据分析平台 大数据开发工程师

2024.05-2024.08

项目描述: 本项目致力于快餐行业线下门店业务的大数据分析与处理。针对该行业复杂的业务流程, 项目采用离线大数据技术构建完善的数据处理体系, 建立集团数据仓库, 对分散在各个门店系统中的业务数据进行整合和预先处理并存储。依据业务分析的具体要求, 从海量的顾客点餐、支付、评价、退款等数据中进行深入挖掘分析, 构建数据集市, 以满足不同业务场景的分析需求。整个项目涵盖多个关键环节: 全量同步使用DataX, 增量同步使用Maxwell、Kafka、Flume构成数据管道, 将全量数据和增量数据传入HDFS并在ODS层同步全部历史数据, 接着传输到DIM层存储维表和DWD层存储事实表, 再传输到DWS层将同一粒度的数据按时间汇总做统计分析, 最后传输到ADS层并将其导出至MySQL, 以供FineBI实现数据的可视化展示, 以便为企业决策提供直观、准确的数据支持。

智能维修数据可视化平台 Java开发工程师

2024.09-2024.12

项目描述: 本项目是为中国烟草公司下属黄鹤楼科技园集团开发的数据统计和分析可视化Web项目。为完善智能维修数据可视化平台的功能, 本项目基于甲方原有系统业务进行需求分析, 先对旧数据表进行重构, 将旧数据经清洗后导入至新数据表, 实现了用户登录、基于条件的下钻统计分析图表可视化、条件查询每年的TOP数据的功能。目前项目已完成上线、交付。Web域名为www.huangheloumaintenancesystem.cn。

## 荣誉奖项

竞赛: 蓝桥杯Java组省二等奖

全国大学生工程训练综合能力竞赛省二等奖

中国工程机器人大赛全国一等奖

中国智能机器人格斗及竞技大赛全国二等奖

证书: 大学英语六级 (CET-6)

软考中级资格-软件设计师