

林靖颐

2000.05 | 女 | (86)15670290309 | linginge@outlook.com

Github: github.com/jylindev



教育背景

香港大学 (硕士)	数据与系统工程	2024.09-2025.12
• 主修课程: 机器学习与应用, 工业工程计算方法, 运筹学, 人工智能在金融领域的应用, 计算机图形学, VR 等。		
郑州大学 (本科)	统计学	2018.09-2022.07
• 主修课程: 统计学, 非参数统计, 时间序列分析, 抽样技术, 多元统计分析原理, 数据挖掘, 大数据分析方法和应用等。		

工作经历

东信时代信息技术有限公司	多模态算法实习生	2025.07-2025.10
• 视频切分: 基于实际应用场景, 构建自造 + 迁移数据集, 并使用 SlowFast (视频帧) 与 HarmonicCNN+MLP (音频帧) 联合建模进行转场分类; 设计标签映射机制, 将原模型 107 类转场标签对齐到火山引擎 API 的 23 类标签, 并引入 “无转场” 新类别以处理误切片段; 最终实现自动定位视频转场帧, 平均准确率由 ~93% 提升至 ~98%。		
• 字幕修复: 基于编辑距离 (Levenshtein Distance) 对齐算法, 实现原剧本文本与自动转写字幕的对齐与修正; 解决自动语音识别 (ASR) 结果中的 错字、漏字、多字、大小写不一致等问题, 提升字幕文本质量。		
• 数字人生成: 基于即梦 API 设计长视频生成 Pipeline, 支持短循环、A-B-A 循环和长视频拓展三种模式, 实现从图片扩展为分钟级视频; FFmpeg 完成帧率检测、首尾帧提取与拼接, 保证视频衔接平滑。		

香港大学人机交互实验室	科研助理 (多模态大模型)	2024.12-2025.06
项目介绍: 基于自构建的 VR 动作视频集处理短视频并生成标注, 用多模态大模型微调以提升动作识别及视频理解准确率。		
• 数据集建立与预处理: 创建并管理自定义 VR 视频数据集, 涵盖了各种人类动作和互动场景 (例如博物馆、画廊、互动空间等); 设计并实施数据收集、标注和预处理流程, 确保数据集包含高质量的图像、视频和对话数据; 结合 GPT 4o 设计三段式模版自动生成的 groundtruth, 为每个视频片段提供精确的动作标签和描述信息, 实现数据集的智能标注和增强。		
• 微调: 以如 LLaVA-NeXT-Video 为生成模型, 训练投影层以及对 llava-next-video 进行 lora 微调。实验表明, 对 VR 动作预测准确率高于基础模型对视频的动作理解, NLG 指标均优于基础模型。		

项目经历

	基于 LangChain 的 Function Call 智能 Agent	2025.05-2025.07
项目背景: 结合国产大模型云雀 (Skylark/豆包) 与 LangChain 框架, 构建支持 Function Call 的智能 Agent, 支持 LLM 在电商与知识检索场景下的工具调用能力。		
技术实现: 接入 Skylark-chat API, 设计 Zero-Shot ReAct Agent 模板。定义多种工具函数 (商品推荐、订单查询、商家政策查询), 由大模型自主选择并调用。构建 ES-knn 向量数据库, 结合 bge-large-zh embedding 实现知识库检索。利用正则与 JSON 解析, 将模型输出转化为标准化 Function Call 请求, 实现 大模型推理 → 工具选择 → 执行 → 答案生成 的闭环。		
成果: 完成可运行的电商场景 Demo (商品导购、订单追踪)。		

CCF-AIOps 挑战赛	基于 Llamaindex+RAG 检索增强的运维知识问答挑战赛	2024.04-2024.07
项目介绍: 面向中兴通讯 5G 网络运维私有化文档 (含 XML/HTML/图像等异构数据) 构建高效 RAG 系统, 采用 CLM-4 作为 LLMs, 通过分析文档的树状结构作为知识路径来优化检索和重排, 在 NLG 指标上(Rouge/Bleu 等)提升了 15%左右。		
• 数据处理: 解析文档目录树生成知识路径, 清洗并转化 HTML 为结构化文本, 使用 glm4v-9b 对图像做 caption 提取图像信息, 最终构建一个带有知识路径、文本和图片信息的多模态知识库; 优化 LlamaIndex 的 SentenceSplitter 分块策略, 减少绝对路径带来的影响, 同时使用关键词对 query 进行改写。		
• RAG 流程: 利用知识路径, 实现常规文本和知识路径双路 BM25 稀疏检索 (文档/路径维度) 和 gte-Qwen2-7B 密集检索, 利用 bge-reranker-v2-m3 重排模型结合知识路径排名, 利用 RRF 算法进行重排, 提高检索内容的可信度,		
• 答案生成: 重排得到的 top6 文本块与图像的内容拼接作为参考内容, 设计多种 prompt 模板得到最优答案。		

技能与语言

熟悉 Python, sql, Shell; 熟悉 linux 开发环境; 雅思 6.5, CET-6 579 分, CATTI 三级笔译; ACCA (英国注册会计师)。