

余坤屹

北京 | 173-3813-3919 (微信同号) | kyu135@ucr.edu | @linkedin | rock3yu.com/zh/

教育经历

加州大学河滨分校 (UCR) 美国加州 GPA 3.88 | 计算机科学 硕士 2024/09-2026/06
课程: 算法, 高性能计算, 机器学习, 人工智能, 编译器, 空间计算, 数据挖掘, 机器人导航, 大语言模型

南方科技大学 (SUSTech) 中国深圳 计算机科学与技术 学士 2020/09-2024/06
课程: 离散数学, 概率论, 算法, 人工智能, 深度学习, 操作系统, 分布式与云计算, 软件工程, 面向对象编程, 数据库等

实习经历

研究生研究助理 | UCR TASL 实验室, 导师: Prof. Jiachen Li 2025/07-2026/06

- 基于 CARLA、SUMO、AirSim 和 ROS 2 开发车辆-无人机协同仿真平台, 并构建基于 7B 视觉-语言-动作大模型的无人机视觉-语言导航 (VLN) 系统
- 构建并发布 SAR-Drone-VLN-3K 数据集, 生成 3000 条搜索-返航轨迹和 6.6 万条图像-动作样本
- 基于 PyTorch 和 LoRA 微调无人机导航策略, 降低验证损失 4 倍、下一步动作预测准确率从 14% 提升至 38%

本科生研究助理 | SUSTech GameAI 实验室, 导师: 刘佳琳、姚新教授 2022/09-2024/06

- 基于 OpenAI 开源多粒子智能体环境, 构建可配置/多环境的追捕博弈仿真平台, 支持多智能体强化学习研究
- 使用 Python 和 PyTorch 实现并训练 MADDPG、QMIX、IQL 等多智能体强化学习算法, 完成 25+ 组实验、累计 5 万轮训练, 证明了 GRU 网络在精度与效率上的权衡, 达到 80%+ 综合追捕成功率

软件开发工程师 | 深圳市蘑菇财富技术有限公司 2023/09-2024/03

- 基于 Vue.js 和 REST API 开发信息中心模块, 服务于 2 家证券公司、500+ 用户
- 集成 37 个 REST API 及 9 个懒加载路由, 通过优化数据获取流程, 将平均页面加载时延从 3.7 秒降至 0.5 秒
- 使用 Feature Flag 扩展配置驱动的权限框架, 实现跨客户的工作流配置, 简化代码分支并提升系统可维护性
- 完成内容数据模型由单选向多选迁移, 覆盖内容创建、审核与发布流程并保持向后兼容; 实现 95% 测试覆盖率并参与 CI/CD

项目经历

ccRAG: 信用卡检索增强问答系统 ccrag.rock3yu.com 大模型 / RAG

- 面向中文用户查询美国信用卡英文条款的场景, 构建覆盖三家银行 54 张卡的中英双语 RAG 系统, 索引 237 份官网页面、条款 PDF 与博客, 全本地部署、零 API 成本
- 系统使用 LlamaIndex 与 BGE-M3 建立 Qdrant 向量索引, 由 Qwen3.5-9B 大模型生成带出处引用的回答
- 自建 122 题人工审核评测集, 按预注册规则完成 50 组对比实验并做配对 bootstrap 检验; 别名过滤与切分重排策略调整后, Recall@5 从 0.62 提升至 0.93, 跨语言召回从 0.34 提升至 0.97, 本地模型答案得分从 0.05 升至 0.85, 与 Claude Sonnet 差距不足 0.1
- 已上线部署于单块消费级显卡, React 与 TypeScript 前端支持流式输出、悬停查看引用原文、拖拽卡片限定检索范围, 并实现请求排队、按 IP 限流与提示词注入防护等安全机制

参观南科大微信小程序 全栈开发

- 开发并部署全栈微信小程序, 服务一万+用户、持续运行半年, 完成了早期访问本校计算机系的管理任务
- 基于 Spring Boot、REST API 和 PostgreSQL 实现角色权限驱动的预约流程, 支持用户提交、管理员审核、查询审批状态及二维码核验等功能
- 使用微信小程序原生框架和 Vue.js 开发前端组件, 并通过 Docker 和腾讯云完成应用容器化部署

技术栈

编程语言 Java, Python, Go (Golang), Swift, C/C++, JavaScript, TypeScript, HTML/CSS

AI 相关 PyTorch, TensorFlow, Transformers/HuggingFace, vLLM, CUDA, Scikit-Learn, LangChain, ROS 2

全栈框架 Spring Boot, React.js, Vue.js, Gin, SwiftUI, Kafka, Redis

数据库/其他 SQL, PostgreSQL, MySQL, GraphQL, REST API, Kubernetes, Docker, 腾讯云, Git, Linux