

杨慎

21岁 | 男 | 个人主页: <https://xi-xiaoran.github.io/>

18248826655 | 2646594598@qq.com



教育背景

南京邮电大学 - 数据科学与大数据技术 - 本科

2022-09 ~ 2026-06

专业成绩: GPA 3.9 (专业前5%)

主修课程: 自然语言处理 (99), 高级程序语言设计 (98), 计算机视觉 (95), 人工智能导论 (95), 算法分析与设计 (94)

实习经验

荣耀 - 算法实习生

2025-05 ~ 至今

- 进行**GEO**(生成式引擎优化), 应用于荣耀400 Pro/Magic V5产品, 在13家媒体投放**67篇**GEO稿件, 实现6小时内14个核心Prompt的**AI可见率 > 70%**。后续持续提供**1000+**篇稿件, 6个月维持AI可见率**> 65%**
- 创新性**地将主流AI引擎的联网搜索推荐机制建模为**二阶段“内容-用户Query”**匹配系统, **逆向分析**其核心特征权重(如**BM25, LDA, 语义相似度**等)基于特征权重分析制定精准GEO策略, 仅优化**5篇**宣传稿, AI可见率从行业底部(约**0.8%**)跃升至**63.5%**(DeepSeek最高达**83.3%**, 其他平台约**55%**), 实现荣耀400 Pro在7个主流关键词搜索结果**全面霸屏**
- 从零到一**独立开发Python自动化脚本, 构建**端到端GEO宣传稿生成链路**, 设计**“反思-迭代”**机制消除语义幻觉, 相比传统品牌方宣传稿, 年化节省成本约**400万元**, 后续优化缩短稿件生成时间**40%**
- 运用**LoRA**与**对比学习**微调DeepSeek-8B模型, 模拟其联网搜索行为进行文献相关性评估, 实现**listwise**级别的先验的文献评估, 构建专业数据集(SEO前45为负样本, DeepSeek联网前5为正样本), 为GEO策略的**先验评估**与优化提供关键数据支撑

博世Bosch - 算法实习生

2024-07 ~ 2024-12

- 基于**Python**对以太网项目历史数据进行清洗和分析, **设计基于规则的自动化决策算法**, 实现网关智能系统覆盖绝大多数场景, 显著降低系统错误率
- 独立实现**Gateway**工具模块的迭代升级, **设计核心解析算法实现自动生成C++通信协议解析代码**。构建自动化测试框架验证生成代码的正确性, 实现从需求文档到可执行代码的全链路自动化
- 独立完成**RTP, RTSP, RTCP**通信协议的**Python**实现与集成, 保障通信链路的稳定性和高效性, 为后续系统开发奠定基础
- 深入优化**Google Protocol Buffer**轻量化库, 完成自动化测试覆盖, 解决边缘场景问题, 提升库的正确率从**98.4%**提高至**100%**
- 利用**Python**可视化工具(如**Matplotlib, Seaborn**)对实际场景数据进行分析, 为性能调优提供数据支撑
- 深度研究**IEEE1588**标准与gPTP协议, 理解其核心算法并完成**Python**实现, 撰写技术文档, 为团队成员提供参考

中国工商银行 - 算法实习生

2024-01 ~ 2024-02

- 熟练运用**Python**编写高效脚本, 处理Excel数据, 完成客户数据的整理、整合与清洗工作, 确保数据的准确性和完整性, 为后续分析提供高质量的数据基础
- 对数据进行前置处理, 包括**数据清洗、特征工程**等, 有效提升模型输入质量, 为后续的数据分析和建模工作提供稳定、可靠的数据支持, 保障模型训练的准确性和稳定性
- 整合来自不同业务系统和外部数据源的客户数据, **解决数据格式不一致的问题**, 建立标准化的数据处理流程, 确保数据的一致性和可用性
- 基于机器学习算法(**KNN, 朴素贝叶斯, MLP**)设计贷款发放智能评估模型, 通过优化算法参数和特征选择, 正确率达到**72%**, 显著提高贷款发放决策的效率和准确性
- 研发**LSTM-Transformer**融合时序预测模型, 在3个月行为数据上达到**89%的AUC**, 提前两周识别**85%高风险客户**

项目经验

开放环境下不确定性感知的可信医学影像分割方法研究(国家级STITP) - 负责人

2024-05 ~ 2025-05

研究背景: 针对现有**证据深度学习**方法在医学图像分割中对边界模糊区域处理不足的问题, 提出了一种创新的**可信医学图像分割**解决方案。

核心贡献: 本人担任**项目负责人**, 完成以下创新点的提出, 所有基于**pytorch**框架代码的编写, 以及所有的实验的设计及实现

- 渐进式证据不确定性引导注意力机制(PEUA): 基于不确定性地图逐步优化模型的注意力分布, 并通过低秩学习降低注意力权重中的噪声, 显著增强困难区域的特征表示能力。
- 语义保留证据学习策略(SAEL): 提出语义平滑证据生成器与保真增强正则化项, 有效避免模糊区域中关键语义信息的丢失, 提升分割的语义一致性。
- 将PEUA和SAEL嵌入至U-KAN框架, **提出 Evidential U-KAN 模型**, 与当前主流方法相比, Evidential U-KAN在准确性和可靠性方面均表现出明显优势。

专利: 《证据不确定性渐进引导的可信医学影像分割方法及设备》

论文: 《Evidential U-KAN for Trustworthy Medical Image Segmentation》(NeurIPS在投, 第一作者, **当前评分5443(6)**, rebutal中)

项目链接: <https://github.com/xi-xiaoran/Evidential-U-KAN>

自我评价

国际化视野: 曾赴**剑桥大学**访学交流, 具备扎实的英语听说读写能力, **雅思成绩 6.5**, 具有良好的国际学术交流能力

初中接触编程, 在**NOIP**大赛中获奖。拥有多年代码开发经验, 具备深厚编程基础。大一获得“蓝桥杯”全国三等奖和美国计算机奥林匹克竞赛(**USACO**)**银奖**

- 大模型: 精通**Transformer, CLIP, BERT**等架构及其应用, 掌握**LoRA**等主流**大模型微调技术**, 并具备利用**强化学习/RAG**优化大模型的实践经验。利用**证据学习**优化大模型, 有效降低模型幻觉。熟悉前沿模型结构如**KAN**网络
- 计算机视觉: 熟练掌握**图像分割/目标检测**等主流框架与方法, 熟悉模型**校准**(Calibration)的主流方法
- 推荐系统: 掌握推荐系统主流算法与架构, 包括但不限于**双塔模型、协同过滤、基于内容的推荐及深度排序模型**, 精通并具备**生成式引擎优化(GEO)**的实践经验
- 不确定性感知: 精通**证据学习**。创新性地提出融合强化学习与元学习的证据学习优化框架, 显著提升了模型预测的准确率与不确定性估计效果, 在相关任务上达到**SOTA**水平。掌握**重尾自正则化**(HT-SR)等**前沿理论**
- 强化学习: 掌握主流强化学习算法(**Reinforce, PPO, DQN, Q-learning**)及其应用, 具备利用RL优化模型(如微调大模型、优化证据学习框架)和解决决策问题的能力