骆明宇

137-6498-9969 luomingyu2002@126.com

教育经历

复旦大学 2025.09 – 至 今
电子信息 直博生 计算与智能创新学院 上海

- 主要研究方向：大模型安全、内生安全、人工智能赋能传统安全

上海海事大学 2021.09 – 2025.06
计算机科学与技术（卓越工程教育） 本科 信息工程学院 上海

- GPA：3.87/4.0，排名:1/31；综合测评：93.85/100，排名:1/31；推免排名：1/108
- 主要课程：数据结构、计算机网络、操作系统、计算机组成原理与汇编、数据库原理与应用等

论文发表

- Multi-Sensor based Strategy Learning with Deep Reinforcement Learning for Unmanned Ground Vehicle, International Journal of Intelligent Networks(EI Compendex, 独立一作)
- An Endogenous Security-Oriented Framework for Cyber Resilience Assessment in Critical Infrastructures. Applied Science. (SCI Compendex, 一作)
- 云边端内核竞态漏洞大模型分析方法研究. 信息网络安全. (CCF-优秀期刊、中文核心，通信作者，导师一作)
- 感知决策系统内生安全韧性度量研究. 中国科学：信息科学. (CCF-顶级期刊，三作)
- RevPRAG: Revealing Poisoning Attacks in Retrieval-Augmented Generation through LLM Activation Analysis. EMNLP. (CCF-B, NLP顶会, 三作)
- On the Security of Tool-Invocation Prompts for LLM-Based Agentic Systems: An Empirical Risk Assessment. S&P. (在投, CCF-A, Security顶会, 共一)
- AutoIPI: Query-agnostic Indirect Prompt Injection on Coding Agents. ACL. (在投，CCF-A，二作)
- Was My Data Used for Training? Membership Inference in Open-Source LLMs via Neural Activations. NDSS. (在投，CCF-A，三作)

实习经历

- **中国银联股份有限公司 金融科技组** 2022.09 – 2024.03
  Python LLM 知识蒸馏 大模型微调 自然语言处理 KIPs
  - 成功部署并微调大语言模型，处理3400+条数据，筛选330+条高质量数据用于反洗钱模型推理，训练loss值稳定在0.46，显著提升模型在可疑行为分析的准确性和可解释性。
  - 开发反洗钱报告自动化评估系统，基于KIPs提取（正则表达式及TF-IDF余弦相似度语义评估）、多维度评估，支持批量处理及可视化图表生成，提升报告生成效率与质量。
  - 搭建Dify平台工作流，解决API调用限制，实现批量数据生成，并优化模型输出筛选流程，确保数据准确性和规整化，显著提升反洗钱报告生成质量。
  - 应用DeepSpeed、LoRA和8-bit量化技术进行模型微调实验，优化qwq3-2B模型性能，针对推理数据不足问题调整微调策略，提升模型输出稳定性。

项目经历

- **基于深度强化学习的无人车运动决策策略学习方法** 论文发表1篇 优秀结题 **负责人** 2022.09 – 2024.03
  Python Pytorch Unity 深度强化学习 注意力机制 多传感器融合 机器人
  - 通过引入**动态多传感器数据融合机制**，提高**多源异构**传感器的数据融合能力
  - 通过在Unity中创建逼真的**虚拟环境**，提高无人车训练速度，实现**虚实迁移**
  - 通过使用**强化学习算法**，使用**经验回放池**等技术，实现无人车**自主避障**

- 与传统方法相比，成功率提高**35.71%**，任务完成时间缩短**37.5%**，训练步数减少**58.33%**

- **自动驾驶车辆交通信号与交警状态监测方法**　负责人　2024.03 – 2024.06
  Python　OpenPose　YOLO　计算机视觉　自动驾驶　多传感器融合
  - 采用**YOLO**算法进行**特种人员识别**，实时识别交通警察和其他相关人员，提高环境感知能力
  - 使用**OpenPose**进行**人体骨架识别**，精确检测和跟踪交警的手势和姿态，识别交通指令
  - 通过**多传感器融合技术**，结合车辆摄像头、雷达和LIDAR数据，构建全方位环境模型
  - 实现**自主避障**功能，在复杂交通环境中自动识别和避开障碍物，同时准确执行交警指令

竞赛及荣誉

中国大学生计算机设计大赛　**全国二等奖**　队长　2023.07
上海市大学生计算机应用能力大赛　**上海市一等奖**　队长　2023.05
蓝桥杯全国软件和信息技术专业人才大赛Python组　**上海市一等奖**　个人　2023.05
中国计算机应用技术大赛-全国算法精英大赛　**华东赛区二等奖**　队长　2024.02
“挑战杯”中国大学生课外学术科技作品竞赛　**上海市三等奖**　队长　2023.05
中国国际“互联网+”大学生创新创业大赛　**上海市铜奖**　队长　2023.07
国家反计算机入侵和防病毒研究中心　信息网络安全专业人员认证——数据安全　2023.03
上海市奖学金(院系4人)、日本邮船奖学金一等奖(院系2人)　2023.09&2022.10
上海海事大学“特等奖学金(院系2%，2次)、优秀团员(2次)、优秀学生(2次)”　2022.10&2023.10

干部经历

计算机类213班、计算机211卓越工程师班　班长　2021.09 – 2025.06
上海海事大学计算机学会　算法部　副部长　2022.09 – 2023.09
上海海事大学计算机学会　会长　2023.09 – 2024.09
中共复旦大学计算与智能创新学院2025级博士生第三党支部委员会　支部纪检委员　2025.09 – 至　今

专业技能

- 掌握Python、C/C++、Java、C#等语言，掌握ROS使用方法，熟练使用Linux与AIGC工具
- 掌握Socket网络编程与爬虫，能够抓取和处理大量网络数据、熟练使用Wireshark进行数据包截取
- 熟悉使用Unity引擎构建虚拟场景，能够使用C#脚本生成Unity动态场景
- 熟练使用SQL与Cypher查询语言，掌握MySQL、Neo4J和Redis数据库使用方法
- 熟悉知识图谱与事理图谱相关知识，了解关系抽取、实体事件对抽取、图异常检测相关内容
- 有一定全栈开发经验，掌握Vue、Vite、Django、Fastapi等框架使用方法，掌握Postman等常用工具使用方法
- 掌握计算机视觉、深度学习与深度强化学习相关方法，掌握OpenCV、YOLO、Pytorch框架使用方法
- 熟悉算法与数据结构，包括排序、动态规划、图论、字符串、数论、数据结构等

其他

- GitHub: https://github.com/kuangren777
- 英语能力: CET-4　CET-6　雅思-6.5