

谢天泓

2453798@tongji.edu.cn | (+86) 183 7203 4526

GitHub: github.com/justimaginary

个人网站: justimaginary.github.io

教育经历

同济大学
信息安全专业 本科

预计 2028 年 6 月毕业
中国 上海

- 均分: 87.4/100 | 核心课程 GPA: 5.0/5.0
- 核心课程: 高级语言程序设计, 数字逻辑, 数据结构, 面向对象程序设计。

研究兴趣

- 自然语言处理、语言模型、结构化预测与可信 AI; 信息抽取、语言模型推理与 AI for Science 场景中的结构化建模方法。

科研经历

滚装码头局地海况要素高分辨率预测模型研究

2025 年 12 月 - 至今

核心成员 | 导师: 李文根 副教授 (先进数据与机器学习系统实验室)

同济大学

- 海量气象数据工程: 基于真实滚装码头及浮标观测数据, 独立设计开发全链路数据预处理管道 (Data Pipeline), 实现多源异构海况与气象特征的自动化清洗、对齐与异常值过滤, 为高精度预测奠定数据基石。
- 模型架构重构与演进: 主导项目核心预测框架 (STD-PLM) 的迭代升级; 基于性能评估结果, 决定逐步淘汰原始基于沙漏注意力 (Sandglass Attention) 的时空编码器, 并引入全新的混合表征学习架构。
- CNN-GCN-LLM 混合预测范式: 创新性地设计并部署基于 CNN (捕获局地微气候空间特征) 与 GCN (建模拓扑网络节点关联) 的联合特征提取器; 探索性接入 LLM (大语言模型) 作为高级推理单元, 增强对复杂海洋气候异常事件的泛化与预测鲁棒性。

基于机器学习与图搜索的大分子材料前体预测

2025 年 9 月 - 至今

第一主持人 | 导师: 马磊 副教授、宋少堂 副教授

同济大学

- 拓扑空间建模: 利用 RDKit 深度解析复杂大分子的原子拓扑, 通过对称环提取 (SSSR) 与 2D 坐标拟合, 将化学结构精准转化为苯环六角网格 (Hexagonal Grid) 逻辑模型。
- 启发式图搜索算法: 设计并实现基于深度优先搜索 (DFS) 的路径规划引擎, 引入动态偏移量学习 (Offset Learning) 与周期性边界条件, 实现高维空间下切分方案的秒级检索与自动化匹配。
- 化学规则引擎集成: 构建智能封端 (Capping) 模块, 严格遵循有机化学合成反应机理, 自动为解构后的骨架添加官能团, 生成具有实际化学意义的异构体前体方案。
- 自动化: 开发全自动处理流, 支持从原始 SMILES 输入到切割方案可视化、前体 SMILES 生成及标准结构图导出的闭环流程, 极大提升了材料设计中单体预测的吞吐量, 并解决了当前化学材料领域中前体分子难以构建的瓶颈问题。

基于底层张量实现的 BiLSTM-CRF 中文分词系统

2026 年 3 月 - 2026 年 4 月

独立开发者 | 技术栈: PyTorch, Python

个人开源项目

- 架构融合: 基于 PyTorch 深度还原 BiLSTM-CRF 机理, 并集成 Transformer 架构中的 8 头自注意力机制与可学习残差缩放因子 γ , 实现局部特征与全局语义的动态加权融合。
- 工程优化: 利用 3D 张量广播与掩码机制彻底重构 CRF 层以实现全并行化计算; 引入 AMP 自动混合精度训练及动态序列截断防御机制, 解决注意力机制在高维场景下的显存瓶颈。

- 性能评估：在 SIGHAN PKU 官方测试集上取得 92.2% 的 F1 分数，并在 OOV（未登录词）比例达 54.5% 的极端测试环境下实现 92.0% 的高召回率，验证了模型极强的泛化与新词发现能力。

基于多尺度深度学习的口腔病灶（溃疡）自动精准识别系统
第一主持人 | 导师：马磊 副教授

2025 年 6 月 - 2025 年 9 月
同济大学

- 医疗数据集构建：主导采集并标注了大量临床口腔溃疡图像，并实现多边形标注向最小外接矩形的精准转换及 COCO/YOLO 格式的跨框架兼容。
- 多范式检测架构对比研究：针对口腔溃疡边界模糊、病灶微小等特征，部署并对比了单阶段 (YOLOv5s/m, YOLOX, RetinaNet) 与双阶段 (Faster R-CNN) 目标检测模型；通过引入 Focal Loss 解决了正负样本极度不平衡的问题，并采用早停策略优化了训练收敛过程。
- 量化性能评估与分析：系统性评估了各模型在测试集上的表现，其中 YOLOv5s 取得了 0.238 的 mAP@0.5:0.95 (多阈值下的平均精度)；同时，RetinaNet 在 AR@0.5:0.95 (平均召回率) 指标上达到 0.338，验证了模型在复杂口腔环境下对微小目标的初步捕获能力。

项目经历

机器学习辅助的团队协作风险预警与防控机制研究
第一主持人 | 导师：余有灵 副教授

2024 年 12 月 - 至今
上海市

- 高并发任务引擎与底层架构：采用 Go 语言结合 Gin 框架构建高性能 RESTful API 集群。针对特定场景设计基于本地 JSON 序列化的轻量级 I/O 存储方案；
- 大模型驱动的智能风险预警闭环：创新性设计结构化多维风险评估矩阵（进度、质量、成员积极性、外部环境）。通过 Go 后端定时守护进程，聚合全生命周期业务数据，对接定制化的 Coze API 智能工作流，实现从底层数据清洗到多源异构数据融合，再到终端生成综合风险等级与应对策略的全链路毫秒级实时预警。
- 基于客观行为画像的量化评价系统：废弃主观 KPI 考核，设计自动化数学建模引擎。通过实时计算成员的积极参与度、执行效率、交付质量与守时性四大核心指标，并在前端通过 ECharts 渲染动态雷达图；将量化指标注入独立的大模型引擎，生成极具参考价值的用人与人力资源调配建议。
- 全栈系统交付与安全底座：前端采用 Vue 3 构建现代化玻璃拟态 UI，集成 Mermaid.js 实现动态拓扑渲染。在网络安全层面，设计并部署基于 JWT 的无状态鉴权机制与细粒度 RBAC 权限隔离中间件，彻底防御越权访问，保障多级审批（双保险屏障）体系下的项目机密数据安全。

QQBotNina: 基于 NapCat 协议的模块化高度可定制机器人
独立开发者 | 技术栈: Python, Docker, SQL

2025 年 10 月 - 至今
个人开源项目

- 异步高并发引擎：基于 Python asyncio 协程架构深度对接 NapCatQQ (OneBot 11) 协议，通过 WebSocket 通信实现非阻塞消息处理管道；在极低内存负载下，确保了大规模并发事件的毫秒级分发与响应。
- 动态插件与解耦设计：设计并实现基于动态导入 (Dynamic Import) 机制的插件框架，支持业务逻辑的热插拔与运行时自愈；通过高度抽象的基类封装指令解析与权限校验层，实现了协议层与业务层的完全解耦。
- 持久化层与数据治理：基于 SQLite 构建高性能异步数据持久化模块，实现对事件日志与用户行为数据的结构化存储；设计复杂查询索引优化策略，为系统的功能迭代与用户画像分析提供数据支撑。
- 容器化部署与稳健运行：构建基于 Docker 的全自动化镜像部署工作流，集成自动化心跳监测与指数退避重连机制，确保系统在校园网复杂拓扑环境下维持 7×24 小时的生产级稳定性。
- 教学应用实践：该系统目前已投入同济大学本科生核心课程《高级语言程序设计》的辅助教学工作中，作为课程自动化管理与交互工具，服务于多名师生。

竞赛与奖项

同济大学本科优秀学生奖学金
学术与综合素质类

2025 年 12 月
同济大学

- 因在学术成绩、科研潜能及综合素质方面表现优异而获此殊荣。