

刘兴岳

大连理工大学 未来技术学院 人工智能专业

aroma@mail.dlut.edu.cn | 15842668147 | recynie.github.io

 Github |  Google Scholar

个人概述

大连理工大学(985)

未来技术学院 人工智能专业

2023 级本科生

学习情况:

- 前五学期学习排名: 14/64
- 大二学习排名: 4/63
- 加权平均成绩: 92.3
- 前五学期综合排名: 22/64
- 大二综合素质排名: 23/263
- 英语四六级: 618, 591

主要专业课成绩:

- 最优化方法: 98
- 计算机视觉: 96
- 计算机视觉基础: 93
- 概率与统计: 96
- 深度学习: 96
- 人工智能程序设计: 92

科研经历

04/2025

现在

自然语言引导的蛋白质优化知识库进化

Overview: 开发了一个用于蛋白质优化的自进化 agent 系统, 通过迭代的“探索-验证-改进”循环提升准确性, 并维护一组代码块作为知识库, 用于积累可执行的先验知识。针对高方差的蛋白质序列空间导致的通用优化策略低效问题, 开发了一个基于代码演化的多智能体系统。系统不再追求单一的全局最优, 而是通过种群演化让每个代码片段学习特定子分布的局部最优, 用多个子分布上的最优解逼近全局最优。

Core Contributions: (1) Code as Knowledge 知识表示: 将结构化的代码种群作为知识库, 每段代码集成门控机制以匹配和优化目标蛋白质分布。(2) 自进化系统: 从头构建高并发的自进化 agent 系统, 将前向探索与知识更新解耦以提升系统效率, 支持多服务器架构提升资源利用率。(3) 搜索算法: 训练时使用 Thompson 采样策略, 推理时采用 LCB score 确定最终知识库, 以此平衡 exploration-exploitation。

Outcome: 项目前期版本获得玻尔+SciMaster AI4S 智能体开发竞赛最佳创意奖 (仅 2 支队伍), 入围奇绩创坛创业营面试 (前 10%)。作为共同第一作者发表 ICLR 2026 RSI Workshop [1]。最新成果预计投稿 NeurIPS 2026。

10/2024

12/2024

分子图表示学习

Overview: 将高层次的分子基序信息分别集成到图神经网络(GNNs)和大语言模型(LLMs)中, 提升分子表示学习能力, 从而突破 1-WL 测试的表达限制以及一维 SMILES 序列表示的局限性。

Core Contributions: (1) Uni-Motif 框架 (GNNs): 设计并验证三种即插即用的融合机制, 将分子子结构信息融入模型; 同时开展对比实验, 评估不同基序切分策略的效果。(2) S^2 Token 策略 (LLMs): 提出并系统评测了一种新的子结构感知分词方法; 复现多个前沿基线模型, 验证基于基序增强方法在性能上的优势及其达到 SOTA 水平。

Outcome: 作为共同作者, 在下游同行评审期刊各发表一篇论文: *Expert Systems with Applications* [2] 和 *Information Processing and Management* [3]。

12/2023

现在

蛋白质水合与结晶条件预测

Overview: 通过建模蛋白质序列、三维结构、化学环境之间的复杂关系, 预测蛋白质结晶条件。

Core Contributions: (1) 数据集构建: 负责完整项目流程, 构建结晶条件数据库 (约 7k 蛋白质), 从 PDB 中提取、筛选、整合结晶条件记录。(2) PLM 微调: 微调 SaProt (预训练的结构感知蛋白质语言模型), 使其能够捕获蛋白质序列、结构、结晶条件之间的潜在关系。(3) 结晶条件相似性分析: 使用 Foldseek 基于结构相似性对蛋白质进行聚类, 分析不同溶液组合的组合效应, 以及相似结构下结晶条件的分布特点。

Outcome: 实验结果表明 PLM 难以有效表征蛋白质结晶条件, 需要引入水合信息等额外信息。目前, 正在尝试把水合信息引入到现有蛋白质嵌入模型中, 构造更强的表征模型。

竞赛奖项

2025/06 | 大连理工大学信息安全竞赛 · 队长

一等奖

2025/09 | 机械工程创新创业赛无损检测 AI 评片赛 · 队长

全国一等奖

奖学金

- 2023-2024 学年, 2024-2025 学年 学习优秀奖奖学金
- 2024-2025 学年 国家奖学金

专业技能

- 深度学习: 自进化智能体, 图神经网络 (GNN), 强化学习, 高参数效率微调 (PEFT), 大语言模型 (LLM)
- AI for Science: 蛋白质表征学习, 分子表征学习, 蛋白质优化
- 程序开发: 编程语言: Python, C/C++; 框架: PyTorch; 开发: Linux, Git, Docker; 写作: LaTeX, Typst

出版物

- [Link] Xingyue. Liu, Zijie. Xing, Runze. Wang, Luoming. Hu, and Yanming. Shen, “Language-guided expertise evolution for protein optimization,” *ICLR 2026 Workshop RSI*.
- [Link] R. Wang, Y. Ma, Xingyue. Liu, Z. Xing, and Y. Shen, “Motif-driven molecular graph representation learning,” *Expert Systems with Applications*, vol. 269, p. 126484, 2025.
- [Link] Runze Wang, Zijie Xing, Xingyue. Liu, Mingqi Yang, Che He, Yanming Shen, “From Graphs to Tokens: Substructure-Aware Molecular Representation for Large Language Models,” *Information Processing and Management*.