个人简历

李旭田



出生年月：2003 年 11 月　　个人主页：lixutian.github.io
手　　机：15512533550　　邮　　箱：a496263365a@163.com
实习意向：代码智能 / 大模型智能体研究实习

教育背景

北京大学 计算机学院 2025.09—预计 2030.06
计算机科学与技术 | 博士（直博） | 软件工程方向

北京理工大学 计算机学院 2021.09—2025.06
计算机科学与技术 | 本科
前三学年平均学分绩 **92.7**，专业排名 **1/162**
获得两次**国家奖学金**、获评北京市三好学生、北京市优秀毕业生等荣誉

研究经历与成果

FeatX | 功能驱动的仓库级代码演化工具 2025.07—2026.05
第一作者 | ASE 2026, Tool and Dataset Track | 已发表：doi.org/10.1145/3832783.3834646

- 针对以代码为中心的交互需手动管理上下文、定位功能实现的问题，以显式功能层次及功能—代码映射驱动代码编辑；结合 JavaParser 静态分析与 LLM 探索式检索构建跨文件上下文，按功能语义变化定位目标、规划修改依赖与执行顺序，输出可审阅的代码差异。
- 在 **38 个真实功能修改提交**上，函数级修改定位 F1 为 **0.385**，较最强对比基线相对提升 **42.6%**；**10 人用户实验**中，平均 NASA-TLX 认知负荷评分较 ChatGPT 对话降低 **41%**。

FeatLens | 功能引导的仓库级代码依赖检索 2025.12—2026.10
第一作者 | FSE 2027 在投 | arXiv 预印本：arxiv.org/abs/2609.26480

- 针对功能需求与分散依赖难以关联、Agent 在线探索需多步 LLM 交互的问题，构建功能—函数索引；沿调用、使用与归属关系扩展实体，融合语义相关性与结构启发式评分，构造个性化 PageRank 重启分布并按预算筛选上下文；离线建索引，**在线依赖检索零 LLM 调用**。
- 相同 Top-15 预算下，DevEval、EvoCodeBench 的依赖召回率较 UniXcoder RAG 提高 **11.2**、**17.4 个百分点**；DevEval 上 GPT-5-mini 的生成 Pass@1 为 **55.03%**。

FeatOutline | 面向代码理解与维护的功能导向仓库摘要 2024.11—2025.11
第二作者 | TOSEM 2026 | 已在线发表：doi.org/10.1145/3862475

- 针对仓库摘要难以兼顾功能组织与细粒度实现关联的问题，形式化功能层次及代码映射，将功能直接关联到实现方法，支持理解与定位；开展文件级、方法级可追溯性评测，区分文件相关性与方法定位精度。在 **32 项真实提交衍生任务**上，方法级 Link@3 F1 为 **57.4%**，最强对比基线为 **13.6%**。

语言模型重复生成分析与神经元干预 2025.09—2026.01
计算语言学课程项目 | 代码：github.com/uniqueFranky/LLMRep

- 复现重复相关神经元定位方法：用 PyTorch forward hook 采集 MLP 激活，按重复前后窗口的平均激活差异筛选候选神经元，并在推理阶段将其激活置零。
- 单独神经元抑制下，GPT-2/IWSLT 的 Rep-3：**70.66%** → **24.54%**；PPL：1.589 → 3.732。

技术能力

编程语言：Java、Python、JavaScript、C++
研究与研发：静态程序分析、代码检索、图算法、LLM Agent 工具实现