

武岳

香港理工大学博士在读

yue0301.wu@connect.polyu.hk | 15333606607 | 个人主页 | Google Scholar | GitHub

个人简介

香港理工大学博士在读，主要研究 LLM 后训练 (SFT、OPD、RL)、高效推理与解码及 LLM Agent Systems，已有多篇第一作者论文被 EMNLP 2026 等会议期刊接收。研究工作涵盖大模型训练算法、Diffusion LLM Decoding、Agent Harness 与 Evaluation，并具有商汤研究院等公司的算法实习经历。

教育背景

香港理工大学	2025 – 至今
博士在读 · 数据科学与人工智能 导师: Jian Huang 教授	中国香港
南方科技大学	2021 – 2025
统计学学士 导师: Bingyi Jing 教授	中国深圳

论文发表与在投成果

† 共同第一作者。

已录用论文

- CoDR: Training-Free Confidence-Drift Remasking for Diffusion Language Models**
Yue Wu, Qinghe Zheng, Yu Zhang, Jian Huang. EMNLP 2026. Accepted. [Paper]
- DSAEval: Evaluating Data Science Agents on a Wide Range of Real-World Data Science Problems**
Maojun Sun[†], Yifei Xie[†], Yue Wu[†], Ruijian Han, Binyan Jiang, Defeng Sun, Yancheng Yuan, Jian Huang. EMNLP 2026. Accepted. [Paper]
- SimLOB: Learning Representations of Limited Order Book for Financial Market Simulation**
Yuanzhe Li[†], Yue Wu[†], Muyao Zhong, Shengcai Liu, Peng Yang. IEEE Transactions on Artificial Intelligence. Accepted. [Paper]

预印本与在审稿件

- Stability-Weighted Decoding for Diffusion Language Models**
Yue Wu, Jian Huang. Under review at ICLR 2027. [Paper]
- Bootstrapped On-Policy Distillation (BOPD)**
Yue Wu, Xueyu Zhou, Yangrong Hu, Jian Huang. Under review at ICLR 2027.
- FamiFT: Familiarity-Adaptive Supervised Fine-Tuning**
Yue Wu, Yangrong Hu, Xueyu Zhou, Jian Huang. Under review at ICLR 2027.
- Self-Privileged Fine-Tuning (SPFT)**
Yangrong Hu, Xueyu Zhou, Yue Wu, Jian Huang. Under review at ICLR 2027.
- DARE: Aligning LLM Agents with the R Statistical Ecosystem via Distribution-Aware Retrieval**
Maojun Sun[†], Yue Wu[†], Yifei Xie[†], Ruijian Han, Binyan Jiang, Defeng Sun, Yancheng Yuan, Jian Huang. Under review at the Journal of the American Statistical Association. [Paper]

科研经历

香港理工大学	2025 – 至今
博士研究	中国香港

- **基于 OPD 的 LLM Self-Improvement:** 提出 BOPD 两阶段框架自蒸馏，利用参考答案监督从初始模型构建任务适应的 Teacher，再通过 OPD 训练独立的 Student，实现无需外部 Teacher 的模型自我提升。在 Qwen3-4B 上，平均数学推理准确率从 18.06% 提升至 23.98%，并超过所构建的 Teacher。
- **自适应 SFT (Adaptive Supervised Fine-Tuning):**
SPFT: 提出基于已有前置信息的 token 加权方法，利用参考答案条件下的冻结模型构建静态监督信号，无需外部 Teacher 或 On-Policy Rollouts。在 Qwen3-8B-Base 上，相较标准 SFT，平均数学推理准确率从 36.41% 提升至 47.80%。
FamiFT: 基于预训练模型对样本的熟悉度，提出自适应 Token 加权方法以调整监督权重。在 Qwen2.5-Math-7B 上，相较标准 SFT，平均数学推理准确率从 21.32% 提升至 37.20%。

科研经历（续） — 香港理工大学

- **Diffusion LLM 高效解码:**
CoDR: 提出基于置信度漂移的免训练重掩码方法，兼容不同解码采样器，对已提交的不可靠 Token 进行纠错。在两个 Diffusion LLM、三种采样器和四项基准上提升准确率；在与 ReMDM 的特定对比设置下，前向计算次数 (NFE) 减少 76%。 [Paper]
SWD: 从信息论角度建立 Token 预测的时序不稳定性与其对掩码上下文的依赖之间的联系，提出稳定性感知解码方法。在 LLaDA-8B 的四项基准上，平均准确率从 25.63% 提升至 41.08%，同时平均 NFE 降低 22.7%。 [Paper]
- **LLM Agents for Data Science:**
DSAEval: 构建面向真实数据科学任务的 Agent Benchmark，涵盖 285 个数据集上的 641 项任务，支持多模态数据、多轮查询及多维度评估，并对 13 个 LLM Agents 进行系统测评。 [Paper]
DARE: 提出 Distribution-Aware Retrieval 方法，将数据统计特征融入 R Package Embedding，在工具检索任务上达到 93.47% NDCG@10（最强 Baseline 为 79.32%），并提升下游 Agent 的统计分析任务表现。 [Paper]

南方科技大学2023.07 – 2024.05

研究助理 | 导师: Peng Yang 教授中国深圳

- **金融市场模拟 (SimLOB):** 开发 Transformer 自编码器，学习限价订单簿 (LOB) 的潜在表示，用于多智能体市场校准。在 10 个模拟场景中的 9 个场景，校准误差低于传统中间价和原始 LOB 方法。 [Paper]

实习经历

商汤研究院 (SenseTime)2025.02 – 2025.05

语音识别与音频数据处理实习生 | 指导: Shiyin Kang

- **GPU 推理优化:** 将原有依赖 CPU 与远程服务器的 ASR Pipeline 迁移至本地 GPU，重构数据加载、批处理及处理模块，使整体处理效率提升 50%。
- **高质量对话数据筛选:** 开发端到端 Data Curation Pipeline，结合 VAD、ASR 及语音质量指标 (SNR、OVRL、停顿时长)，从大规模对话数据集中识别并提取高质量对话。
- **命名实体识别 (NER):** 训练基于 BERT 的 NER 模型，从音乐文件名中提取歌手和歌曲名称，结合 Embedding 相似度匹配进行重复数据检测，并构建结构化元数据。

xDAN AI2024.03 – 2024.06

实习团队负责人 | 指导: Yipeng Cheng

- **合成数据生成 (Synthetic Data Generation):** 开发针对于 LLM 指定能力的数据增广流程，覆盖 10 余个数据集，包括诗歌等领域任务，提升指令微调数据的多样性与覆盖范围。
- **数据质量评估:** 借鉴 DEITA 和 UltraFeedback，开发基于 LLM 的数据评分框架，评估样本复杂度与质量，支持高质量指令微调数据集的构建。

专业技能

- **编程:** Python (PyTorch、NumPy、pandas、scikit-learn)、Java、C++、 $\LaTeX$
- **开发工具:** Git、Docker、Linux
- **语言能力:** 普通话 (母语)，英语 (流利; TOEFL: 103)

获奖荣誉

- 南方科技大学优秀毕业生2025
- 南方科技大学校训“创新”系列奖学金特等奖2024
- 南方科技大学优秀学生奖学金一等奖2023, 2024
- iGEM 国际竞赛金奖2023
- 全国大学生数学建模竞赛广东赛区一等奖2023