

江宸宇

手机号码: (+852) 55711843

电子邮箱: jchenyu@connect.hku.hk

个人主页: <https://chenyu-jiang.github.io/>

微信: 18707122931

地址: 412 室, 周亦卿楼, 香港大学, 香港特别行政区

教育背景

香港大学 **The University of Hong Kong (HKU)** 2020 年 9 月-现在 (2025 年底毕业)

- 学位: 博士, 计算机科学 (Ph.D. in Computer Science)
- 导师: 吴川教授 (Prof. Chuan Wu)

香港中文大学 **The Chinese University of Hong Kong (CUHK)** 2016 年 9 月- 2020 年 8 月

- 学位: 学士, 计算机科学 (B.Sc. in Computer Science)
- 以甲等荣誉 (**First Class Honors**) 毕业

研究方向

主要研究神经网络/大模型的训练系统优化。主要课题包括动态输入场景下上下文及流水线并行优化, MoE 模型训练优化等。其他研究兴趣也包括分布式训练/推理中网络优化和深度学习编译器等等。

论文发表

-
- **Chenyu Jiang, Zhenkun Cai, Ye Tian, Zhen Jia, Yida Wang, Chuan Wu, DCP: Addressing Input Dynamism in Long-Context Training via Dynamic Context Parallelism, in SOSP'25**
 - o 主题: 在动态输入 (序列长度, Attention Mask) 场景下提高上下文并行训练效率。
 - **Chenyu Jiang, Zhen Jia, Yida Wang, Shuai Zheng, Chuan Wu, DynaPipe: Optimizing Multi-task Training through Dynamic Pipelines, in EuroSys'24**
 - o 主题: 在动态序列长度场景下优化 micro-batch 切分及流水线调度提升训练效率。
 - **Chenyu Jiang, Ye Tian, Zhen Jia, Chuan Wu, Yida Wang, Shuai Zheng, Lancet: Accelerating Mixture-of-Experts Training by Overlapping Weight Gradient Computation and All-to-All Communication, in MLSys'24**
 - o 主题: 通过调度权重梯度计算以及增大模型切分范围提高通信计算重叠。
 - Hanpeng Hu, **Chenyu Jiang**, Yuchen Zhong, Yanghua Peng, Chuan Wu, Yibo Zhu, Haibin Lin, *dPRO: A Generic Performance Diagnosis and Optimization Toolkit for Expediting Distributed DNN Training*, in **MLSys'22**
 - o 主题: 收集分布式训练中各机计算及通信轨迹, 整理全局的数据流图以供分析, 预测及优化。
 - Shiyuan Deng, Xiao Yan, Kelvin K.W. Ng, **Chenyu Jiang**, James Cheng, *Pyramid-A General Framework for Distributed Similarity Search*, in IEEE Big Data 2019 (Short Paper)

实习经历

美团 2025 年 5 月-现在

- 团队: 大模型预训练

亚马逊云科技 **Amazon Web Services** 2021 年 6 月-2025 年 2 月

- 职位: Applied Scientist Intern

演讲

- “Optimizing Distributed Large Language Model Training Under Novel Parallelisms and Dynamic Inputs”, Anuttacon, 线上, 2025 年 10 月
- “DCP: Addressing Input Dynamism in Long-Context Training via Dynamic Context Parallelism”, SOSPP’25 会议, 首尔, 韩国, 2025 年 10 月
- “Lancet: Accelerating Mixture-of-Experts Training by Overlapping Weight Gradient Computation and All-to-All Communication”, MLSys’24 会议, 圣何塞, 美国, 2024 年 5 月
- “Optimizing novel DNN training scenarios-from data to model architecture”, Bytedance Ltd, 圣何塞, 美国, 2024 年 5 月
- “Optimizing novel DNN training scenarios-from data to model architecture”, Amazon Web Services (AWS), 圣何塞, 美国, 2024 年 5 月
- “DynaPipe: Optimizing Multi-task Training through Dynamic Pipelines”, EuroSys’24 会议, 雅典, 希腊, 2024 年 4 月

获奖情况

- HKU University Postgraduate Fellowship [2020]
- Professor Charles K. Kao Research Exchange Scholarship [2019]
- ELITE Stream Student Scholarship [2017-2019]
- CSE Outstanding Academic Performance [2018-2019]
- Wu Yee Sun College Academic Excellence Scholarship [2017-2018]

教学经验

- 教学助理, COMP9602 Convex Optimization (凸优化), 香港大学
- 教学助理, COMP3322 Modern Technologies on World Wide Web (万维网现代技术), 香港大学
- 教学助理, COMP3234 Computer and Communication Networks (计算机网络), 香港大学

其他研究经历

- 毕业设计: 分布式深度学习系统, 香港中文大学 2019 年 9 月-2020 年 8 月
导师: Prof. James Cheng
描述: 研究数据并行下 DNN 训练中通信的加速。
- 暑期实习: 机器学习中的公平性问题, 美国卡耐基梅隆大学 2019 年 6 月-2019 年 8 月
导师: Prof. Pulkit Grover
描述: 研究用信息论工具减少推荐系统中系统性的偏见与歧视。
- 研究实习: 自然语言处理中的表征学习, 香港中文大学 2018 年 6 月-2019 年 1 月
导师: Prof. Helen Meng, Dr. Kerney Wu
描述: 探索基于字典数据生成多词义词嵌入的方法。