

陈江海

(+86) 186-7172-0497 · cjh18671720497@outlook.com · CjhHa1.github.io · MLsys

教育背景

香港大学, 人工智能, 硕士在读

2024.9-2026.1(预计)

QS TOP20

- 课程：人工智能基础、人工智能优化、应用数据挖掘与文本分析

南开大学, 智能科学与技术, 工学学士

2019.8-2023.6

GPA: 3.67/4.0 (87.49/100),

- 课程：高级编程语言 (93.4), 概率论与数理统计 (92), 微积分 (94), 机器学习 (94), 计算机视觉 (92), 深度学习 (94), 强化学习 (95)

所获奖项

- 国家励志奖学金 (两次)
- 美国大学生数学建模竞赛 (MCM) 荣誉奖

研究经历

新加坡国立大学, 高性能计算与人工智能实验室 & 华为 | 诺亚方舟实验室

2022.3-2023.6

研究实习生。负责扩展 *Transformer* 的高效训练框架。

- 项目：研究 *Transformer* 配置与训练目标之间的关系。由总统青年教授尤洋指导。(ICML 2023)
- 职责：收集预训练数据，进行实验，执行理论推导。

南开大学, 机器人与自动信息系统研究所

2021.9-2022.6

学生研究实习生。负责用于故障诊断的对比学习方法。

- 项目：研究通过自监督预训练 (对比学习) 方法解决数据稀缺场景下的轴承故障诊断问题。由杨博远教授指导。(ISIE 2022)
- 职责：部署 Momentum Contrast (MOCO) 预训练框架用于轴承故障诊断模型，缓解现有训练模式在有限标记样本场景下面临的困难。

国家级大学生创新创业训练计划

2021.4-2022.4

- 项目：基于机器学习的台风灾害快速损失评估与救灾应用：以宁波为例。项目由陈晓伟教授指导。
- 职责：开发机器学习方法，选择致灾指标，进行实验分析。撰写研究论文一篇，项目获天津市二等奖。

工作经历

百度 | 深度学习技术平台部, DLTP

2022.10-2023.4

AI 实习生。负责大规模神经网络的自动并行算法。

- 自动并行：协助设计和实现了针对 *PaddlePaddle* 的基于规则的全自动并行算法。创新地将模式匹配策略引入到两阶段搜索算法中，显著提高了搜索效率。
- 项目贡献：参与算法设计、算子开发，并对包括 *GPT* 和 *ResNet* 在内的多种模型进行了实际验证。负责撰写研究论文。

潞晨科技 | HPC-AI LAB, System

2023.6-2024.7

机器学习系统工程师。负责深度学习系统和高性能推理框架。

- 训练系统：Colossal AI 流水线并行架构设计，用于部署大规模神经网络的分布式系统。开发基础组件并进行了实际测试。
- 推理框架：设计及开发了 Colossal-Inference，一个高效且轻量级的推理框架。
- 自动并行：设计并开发了一种基于性能分析的自动并行方法，用于国产芯片大规模训练的分布式算法自动调优。

华为 | 诺亚方舟实验室 (香港), 语音语义

2024.11-2025.7

研究实习生。负责使用确定性路由训练和优化专家混合模型 (MOEs)

- 路由策略：探索用于专家混合模型 (MOE) 训练的确定性路由策略及其他策略。
- 升级训练：研究将稠密模型继续训练为专家混合模型 (MOE) 的升级训练方法 (upcycling)。

论文

- **A Study on Transformer Configuration and Training Objective:** Fuzhao Xue, Jianghai Chen, Aixin Sun, Xiaozhe Ren, Zangwei Zheng, Xiaoxin He, Yongming Chen, Xin Jiang, Yang You. ICML 2023
- **Self supervised Contrastive Learning Approach for Bearing Fault Diagnosis with Rare Labeled Data:** Jianghai Chen, Boyuan Yang, Ruonan Liu ISIE 2022

技能 & 自我评估

- **语言能力：**母语为普通话，英语流利（雅思 7.0，单项不低于 6.0），持有 CATTI 三级证书
- **软件技能：**精通 MS Office 和 LaTeX，掌握多种编程语言，如 C++、Python 等
- **自我评价：**目标导向，充满热情和意志力，理性且逻辑性强，具有求知欲。