

孙平炜

☎ (+86)18641202035 | ✉ 1738860308@qq.com | 🏠 polarispw.github.io/ | 🌐 github.com/polarispw

教育背景

香港科技大学

中国, 香港

Master of Science in Big Data Technology

2023 年 8 月 - 2025 年 1 月

- **Topics of interest:** Efficient LLM, Compression and acceleration of LLM

东北大学

中国, 辽宁

人工智能-工学学士

2019 年 9 月 - 2023 年 7 月

- **GPA:** 90.23/100
- **修读课程:** 概率论, 离散数学, 最优化方法, 数据结构与算法, 计算机系统, 模式识别, 机器学习, 深度学习, 自然语言处理
- **课外活动:** 东北大学党委宣传部视觉工作室-负责人

专业技能

编程语言 Python, C/C++, Verilog

框架 & 工具 Pytorch, Deepspeed, MNN, TNN, llama.cpp, Git, Docker, Huggingface, Vivado

实践经历

iSING Lab, 香港科技大学

中国, 香港

Research Intern

2023 年 9 月 - 至今

- **Research topic:** AWSVD: Compressing LLM with Weight and Activation weighted SVD (competitive with LLM-Pruner and SliceGPT)

OPPO 研究院

中国, 北京

算法工程师 (实习)

2023 年 3 月 - 2023 年 7 月

- **NLP 模型高效部署:** 改进 BERT 类模型在智能手机上的推理性能, 调研 LLM 在智能设备上部署的技术方案 (已发布于 ColorOS)
- **高性能矩阵乘算子:** 为自研的大模型推理框架移植 GEMM 算子 (ARM 平台)

计算机科学与工程学院, 东北大学

中国, 辽宁

计算机系统-课程助教

2022 年 9 月 - 2023 年 1 月

- **工作内容:** 实验室环境配置、课程讲解以及项目验收评分。

个人项目

Fine-tuning vs Prompting, Can Language Models Understand Human Values?

中国, 香港

計算機科學與工程系, 香港科技大學

2023 年 9 月 - 2023 年 11 月

- 使用人类价值观数据集对比了下游任务微调和端到端微调 (提示微调) 两种方法, 通过不同结构不同尺度的模型对比检验预训练模型对人类价值观的理解能力。

基于 MIPS 指令集的高性能 CPU 设计及 AI 应用

中国, 辽宁

计算机科学与工程学院, 东北大学

2023 年 1 月 - 2023 年 6 月

- 使用 Verilog 在 FPGA 开发板上搭建基于 MIPS 的双发六级流水线 CPU, 配合 CNN 加速核心实现对 MINIST 数据集的硬件级加速处理。

基于 BERT 模型的情感分析

中国, 辽宁

NEU NLP-Lab

2022 年 6 月 - 2022 年 8 月

- 该项目基于 BERT 模型, 通过多种微调方法和对比学习任务来增强模型在某一极度不平衡数据集上的分类性能。

曾获荣誉

2020 国家奖学金, 中华人民共和国教育部

中国, 辽宁

2020 一等奖学金, 东北大学

中国, 辽宁

2021 二等奖学金, 东北大学

中国, 辽宁

2022 “龙芯杯”全国三等奖, 全国大学生计算机系统能力培养大赛组委会

中国, 北京