# kozumi

\_\_\_\_\_\_\_ | ≥ example@email.com | 🞧 yourusername | 广东省广州市

# **命**教育背景

华南理工大学 2024.09 - 2028.06 (预计)

本科, 机械电子工程 | GPA: 待补充

- 核心课程成绩(教学班第1名):
  - ▶ 线性代数与解析几何(97)
  - ► C++程序设计基础 (97)
  - ▶ 学术英语 (91)
- 英语能力: CET-4 632 (2024年12月)

## 亟 技术能力

编程语言 Python (熟练), C++(熟练), Rust (项目实践)

开发工具链 ▶ 版本控制: Git/GitHub

► 构建工具: CMake

▶ 容器化: Docker

► 包管理: conda/pip/pipenv ▶ 开发环境: VS Code/Bash 脚本

深度学习 熟悉大模型推理优化技术(vLLM/TGI等),了解量化部署方案(AWQ/GPTQ等)

## ♥ 项目经历

大模型推理系统开发 202X.XX - 202X.XX Rust/CUDA/Python

InfiniLM 训练营项目

- 基于 Rust 实现 Transformer 架构推理系统, 支持动态张量计算与注意力机制
- 完成 CUDA 内核优化, 实现比原生 Python 实现快 3 倍的推理速度
- 开发配套 C++ 接口模块, 通过 SWIG 实现 Python 绑定
- GitHub 仓库: yourusername/learning-lm-rs

#### 开源社区贡献

- PaddlePaddle 社区:提交 X 个 PR(具体说明模块优化/功能开发内容)
- GIS 文档翻译:完成 XX 万字技术文档中英翻译(建议补充具体模块名称)

### ♥ 竞赛实践

#### 紫光同创杯 FPGA 设计大赛

202X.XX - 202X.XX

- 基于 Vivado 实现 XXXX 功能模块(建议补充具体技术细节)
- 完成 RISC-V 处理器与 DDR3 控制器的接口设计

# 赛元杯电子创新设计大赛

202X.XX

• 开发基于 STM32 的 XXXX 系统(建议补充传感器/算法应用)

#### "机甲杯"机器人设计大赛

202X.XX

• (建议补充具体技术实现)

## ❷ 技术视野

持续学习 通过技术博客/论文阅读跟踪 LLM 最新进展

工具链 熟悉大模型开发完整工具链(训练/推理/部署/监控)

# ● 补充说明

- 建议补充 GitHub 链接展示代码能力
- 建议补充具体项目指标(如模型参数量、推理速度提升百分比等)
- 建议补充实验室科研经历(如有)