

陈洪熠

+86 186 5920 6628 | tomchen0417@gmail.com

教育经历

香港中文大学（深圳） 数据科学 本科 2023.09 - 至今

- GPA: 3.5/4.0 (专业前 25%)
- 核心课程: 机器学习、最优化、随机过程、离散数学、常微分方程、概率统计、线性代数、数据结构等。
- 荣誉奖项: 院长荣誉录 Dean's List、知行社会实践二等奖。

加州大学伯克利分校 (UC Berkeley) 计算机科学 访问学生 2025.08 - 2025.12

- 核心课程: 人工智能、计算机体系结构、深度神经网络。

实习经历

杉数科技 算法实习生 - 机器学习方向

上海 2026.07 - 至今

招商银行 金融科技实习生

福建, 厦门 2025.07 - 2025.08

- 模型开发: 参与基于 BERT 的反欺诈文本分类项目, 面向数万条多源金融文本构建三级风险分类; 通过模型微调、超参数搜索及高风险预警阈值优化, 使准确率较 XGBoost 基线相对提升 12%。
- 可解释性与落地: 基于 SHAP 聚合分析 Token 级特征贡献, 识别高风险语义模式并支持错误分析与阈值优化; 模型已部署至银行内部测试环境。

数超科技 数据分析实习生

广东, 深圳 2025.04 - 2025.06

- 特征分析: 使用 Python/SQL 对广告投放数据进行处理与分析, 识别影响转化效率的关键因素, 为投放策略与产品运营决策提供量化依据, 推动核心产品竞价成功率提升 20%。
- 策略优化: 在 RTB 场景下设计 A/B 测试方案, 对竞价策略进行对比验证并优化, 使 ROAS 提升 10%。

科研经历

CardioFormer 心源性猝死预测

导师: Prof. Joshua Zhou 2026.04 - 至今

- 患者状态建模: 基于 EHRFormer 编码纵向 EHR 中的离散、连续和时序信息, 构建动态患者状态空间表示, 用于高风险心血管事件预测。
- 主动检查策略: 构建 IQL 强化学习策略, 在历史检查轨迹上联合建模信息增益、检查成本与等待风险, 自适应选择下一项检查或 STOP, 最大化检查效益。
- 安全决策机制: 引入历史医生策略约束及 Safety Critic + STOP Shield, 避免偏离临床行为的检查建议与高风险患者过早停止, 在保持高风险识别能力的同时复合优化检查数量、成本与决策时间。

SkinGPT-X 医学大模型

导师: Prof. Joshua Zhou 2025.12 - 2026.04

- 评测体系: 参与构建医学大模型评测框架, 覆盖诊断准确性与复杂病例处理能力等核心维度。
- 记忆模块: 评测 Qwen3、MedGemma、HuLuMed 等 SOTA 模型, 并识别关键错误模式, 进而设计 Self-Evolving Memory。
- 项目成果: SkinGPT-X 在多项指标上实现约 8% 的性能提升, 并在复杂病例任务中展现出更高的鲁棒性; 相关成果已于 2026 年 3 月投稿至 Nature Medicine (审核中)。

多语义性对大模型编辑 (Model Editing) 的影响

导师: Prof. Anant Sahai

[GitHub Repo](#)

2025.11 - 2025.12

- 实验设计与验证: 独立设计并验证核心实验, 研究大模型中“多语义神经元”在知识编辑任务中的影响机制, 证明神经元级编辑方案 (FiNE) 相比层级编辑方案 (ROME) 更易引发附带损害。
- 框架搭建与分析: 在 Pythia-6.9B 模型上独立搭建完整验证框架, 综合运用因果追踪与碰撞分析方法定位多语义神经元, 量化对比附带损害, 并分析 FiNE 与 ROME 在精确度与鲁棒性之间的权衡关系。

项目经历

中远海科校企合作

项目负责人

2026.07 - 至今

- 航贸宝: 主导面向大宗商品交易的基本面分析平台建设, 基于 FastAPI、PostgreSQL 打通船舶、港口、运价与到港等多源数据; 构建 LangGraph + Function Calling 多步骤 Agent, 将智能问数、舆情分析与异常提醒集成至统一分析链路。
- 航运港口智能问数系统: 设计 NL2SQL Agent, 结合业务词典、实体规范化、澄清追问与参数校验, 将自然语言问题转化为可执行 SQL; 融合 AIS、雷达、报批、执法、港口及 S-100 地理时空数据, 自动生成查询结果、图表与趋势分析。

Horizon: AI 个性化新闻雷达 (开源项目, 7k+ Stars)

[GitHub Repo](#)

开源项目贡献者

2026.06

- Agent 工作流: 基于 Python 异步任务构建端到端信息处理链路, 串联多源抓取、内容去重、LLM 重要性评分、摘要生成与 Webhook 自动分发。
- 筛选策略: 设计 Balanced Digest 机制, 结合 LLM 评分与类别配额筛选内容, 提升主题覆盖度并减少无效调用。

Ranno.ai (Google AI 黑客松)

[GitHub Repo](#)

核心开发者

2025.10

- 端侧 AI 架构: 基于 Chrome Prompt API (Gemini Nano) 实现端侧自然语言理解模块, 将用户输入解析为结构化意图。
- 路线规划系统: 融合 Google Maps API、犯罪率与城市设施数据, 使用 React、TypeScript 构建混合路线规划与交互系统。

技能专长

编程语言 & 前端开发

Python, R, SQL, Java, C, JavaScript, HTML/CSS, LaTeX, Git

机器学习 & 数据分析

PyTorch, Scikit-learn, Transformers, Pandas, NumPy, Tableau

语言能力

普通话 (母语), 英语 (托福 109, 能作为工作语言)

课外活动经历

知行社会实践项目

项目负责人

2024.06 - 2024.10

- 负责数据处理与分析, 使用 Python (Pandas) 完成问卷数据清洗与特征构建, 并基于 K-Means 对反馈进行聚类, 提取 4 类核心用户画像。
- 输出数据分析报告, 核心结论被香港中文大学 (深圳) 采纳, 用于学生工作策略优化。

乡村振兴三下乡项目

联合项目负责人

2024.08

- 负责多主体调研设计与执行, 带队完成政府与村民访谈及一手数据采集。
- 将访谈内容结构化为文本数据, 基于主题分析提取乡村产业发展关键问题与政策反馈。
- 撰写调研报告, 被地方政府采纳并获当地媒体报道。