

孟宏道

+86-18500305220 | mycrofthd@gmail.com | [LinkedIn](#) | [GitHub](#) | menghongdao.com

教育背景

纽约大学

理学硕士学位, 计算机科学与技术 | GPA: 3.8

纽约, 美国

2024.9 - 2026.5 (预计)

北京工业大学

工学学士学位, 信息安全 | GPA: 3.6

北京, 中国

2020.9 - 2024.7

技术栈

- 编程语言:** Java, Python, C/C++, Go, SQL, JavaScript, TypeScript, HTML/CSS, Shell, PHP, Latex
- 架构:** React, Angular, Vue.js, Django, Flask, Node.js, Spring Boot, PyTorch, Tensorflow, Pandas, Scikit-Learn
- 数据库:** MySQL, Redis, MongoDB, PostgreSQL, DynamoDB, Oracle, Firebase, RocketMQ, Elasticsearch, MilvusDB
- 其他工具:** Git, Docker, AWS, Azure, CMake, Postman, CI/CD, Jenkins, Nginx, LangChain, Jira, Figma, Ali Cloud

工作经历

C2SMARTER Center | 机器学习工程实习生 | 纽约, 美国

2025.1 - 现在

- 基于 **Django + Kafka** 构建实时物联网分析管道, 每日处理 **720k+** 传感器事件 (约 83.3 条/s), 通过时间窗口聚合实现 P95 端到端延迟 $\leq 80ms$, 并准确触发预测性维护提醒
- 部署 **Nginx** 反向代理集群, 通过动态负载均衡与 SSL 终端卸载优化请求路由, 将服务器响应时间降低 **23.7%**
- 运用 **Terraform** 进行 AWS 三区域基础设施即代码部署, 使环境配置时间由 4 小时缩短至 50 分钟 (效率+79%)
- 通过 **Ansible Playbook** 定义滚动更新策略, 将部署失败率降低 **67%**, 版本回滚时间缩短至 ≤ 3 分钟
- 设计跨 3 可用区的 **Kafka** 集群架构 (因子=3), 通过监控驱动自动分区再平衡, 实现 99.95% 可用性

深度伪造检测 (初创公司) | 机器学习工程师, 创始人 | 纽约, 美国

2024.9 - 2024.12

- 基于 **React + TypeScript** 开发深度伪造检测平台前端核心模块, 实现实时交互并支持 **1200+** 并发用户, 客户端渲染延迟降低 21.3%。采用 **Django + WebSocket** 构建实时通信层, 通过长连接复用与消息压缩协议保障中间件通信 P99 延迟 $< 180ms$, 整体任务处理效率提升 **25.6%**
- 在 **AWS EKS** 托管 **Kubernetes** 集群上部署推断服务, 结合 **AWS ELB** 流量分发与 **HPA** 机制, 在 5k RPM 负载下实现 99.5% 系统可用性
- 通过 **EXPLAIN ANALYZE** 分析 **PostgreSQL** 执行计划并优化复合索引, 将平均响应时间从 320ms 降至 262ms
- 基于 **Hugging Face** 的 **Vision Transformer** 与 **VGG16** 模型进行图像检测微调, 利用 **OpenCV** 和 **FFmpeg** 从 **FaceForensics** 视频中提取帧与音频特征; 图像检测准确率 **91.2%**, 音频检测准确率 **88.1%**, 并将模型部署至 **AWS EC2**
- 作为 **Scrum Master** 主导 5 人跨职能团队 (前端/后端/QA), 通过 **Jira** 需求拆解与 **Figma** 高保真原型驱动开发, 使需求交付准时率提升 20%

青藤云安全, 产品研发部门后端组 | 机器学习工程实习生 | 北京, 中国

2024.2 - 2024.7

- 开发全栈 **RAG** 智能问答系统: 基于 **React + Redux** 的响应式前端与 **Flask** 后端, 集成 **LangChain** 框架优化检索流程, 用户参与度提升 **33.7%**, 数据检索延迟降低 **23.6%**。
- 采用 **MongoDB** 分片集群 + **Milvus** 向量数据库 (HNSW 索引) 构建混合存储方案, 并利用 **Docker** 容器化部署, 查询性能提升 **25.6%**, 部署效率提高 **15.7%**
- 通过 **Jenkins** 多分支流水线集成 **SonarQube** 代码检测, 实现 **CI/CD** 自动化, 部署周期缩短 **18.7%**, 系统成功支持并发用户从 5K 增长到 10K
- 基于 **Prometheus + Grafana** 搭建可观测性监控体系, 定义 P95 延迟/错误率等 12 项核心指标, 实施日志链路追踪与告警优化, **MTTR** 降低 **25.3%**
- 制定 **GitLab Flow** 分支管理规范与 **Postman** 自动化测试集, 解决 82% 的集成冲突问题, 团队协作效率提升 22%。

研究经历

数据挖掘与安全实验室, 北京工业大学 | 机器学习工程实习生 | 北京, 中国

2022.9 - 2024.7

- 主导了联邦学习和多视图多标签机器学习的研究, 重点关注隐私保护下的特征融合和多视图多标签分类。在 **IEEE Transactions on Big Data 2025** 上发表了第一作者论文 **《Federated Multi-View Multi-Label Classification》**
- 提出了并开发了 **FMVML** 框架, 支持跨视图特征融合和多标签语义分类, 成功解决了复杂的数据隐私挑战, 超越了当时所有最先进的方法, 提高了 **8.3%** 的平均精度, 并降低了 **14%** 的 One Error 指标。
- 使用 **Python/PyTorch** 和 **Matlab** 进行模型开发; 利用 **Pandas/Scikit-Learn** 实现数据管道; 通过 **LATEX** 撰写论文。

论文发表

- 第一作者: "Federated Multi-View Multi-Label Classification." IEEE Transactions on Big Data, 2025.
- 共同作者: "Susceptibility genes identification and risk evaluation model construction by transcriptome-wide association analysis for salt sensitivity of blood pressure: the EpiSS study." BMC Genomics, 2024.
- 第一作者: "基于数据隐私保护的多模态信息融合关键方法。" 国家知识产权局 (发明专利, 申请号: 2023111952962)