

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:** C/C++, Java, Python, Go, SQL, JavaScript, TypeScript, HTML/CSS, Shell, PHP, Latex
- **架构:** React, Angular, Vue.js, Django, Flask, Node.js, Spring Boot
- **数据库:** MySQL, Redis, MongoDB, PostgreSQL, DynamoDB, Oracle, Firebase, RocketMQ, Elasticsearch, MilvusDB
- **机器学习:** PyTorch, LangChain, Tensorflow, Pandas, Scikit-Learn, VGG16, Vision Transformer, Multi-View, Multi-Label
- **其他工具:** Git, Ali Cloud, Docker, AWS, Azure, CMake, Postman, CI/CD, Jenkins, Nginx, Jira, Figma

工作经历

字节跳动 | 全栈开发实习生 | 西雅图, 美国

2025.5 - 2025.9

- 开发基于 Python 的 DAG-DSL 迁移工具包 (结合 Protobuf 与 regex), 实现字节跳动特征服务 (BFS) 老版本配置文件到标准化 DSL 层的自动转换, 使算子迁移吞吐量提升 85%。
- 设计并实现基于 Blade 的 C++ 算子迁移 pipeline, 显著优化 BFS 迁移流程, 算子迁移速度提升 83.3%, 并个人完成 20 个算子的迁移。
- 设计并开发 User Data Accessor (UDA) 的一站式 workflow (推荐引擎的数据提取服务), 涵盖服务构建编译、Python 算子自动生成、DSL 编译与 RPC 请求, 支持 40 多名工程师参与 UDA 的开发和测试, 开发效率提升 73.3%。
- 构建基于 DAG 的服务引擎 C++ 调试框架, 利用 gflags 控制按需日志记录, 实现拓扑有序子图内节点执行细节的追踪。

C2SMARTER Center | 机器学习算法工程实习生 | 纽约, 美国

2025.1 - 现在

- 基于 React+Redux 前端与 Flask 后端开发 RAG (检索增强生成) 聊天机器人系统, 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%
- 使用 MongoDB 分片集群+Milvus 向量数据库与 RESTful API 构建可扩展的后端架构, 结合 Docker 容器化部署, 查询性能提升 25.6%, 部署效率提高 15.7%
- 在 AWS EC2 上部署应用并通过 Jenkins 实现 CI/CD 自动化测试与高可用部署, 支持并发用户数从 5k 提升至 10k。
- 可观测性监控体系搭建: 基于 Prometheus+Grafana 定义 12 项核心监控指标 (P95 延迟/错误率), 实现日志链路追踪与告警优化, MTTR 降低 25.3%

Freeleaps | 全栈开发实习生 | 西雅图, 美国

2025.3 - 2025.5

- 设计并开发基于 AI 的 CI/CD 自动化平台, 集成 GitHub Actions、Docker、Jenkins 与 Kubernetes, 实现智能构建优化与零停机部署。
- 使用 TensorFlow、LangChain 与 GPT-4o API 构建智能日志分析与故障预测模块, 加速构建恢复流程, 整体部署时长缩短超过 40%。

深度伪造检测 (初创公司) | 机器学习算法工程师, 创始人 | 纽约, 美国

2024.9 - 2024.12

- 主导设计基于 React+TypeScript 的深度伪造检测平台核心模块, 实现了实时交互并支持 1200+并发用户。
- 使用 Django+WebSocket 构建实时通信中间层, 保障中间件通信 P99 延迟在 180ms。
- 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
- 基于 AWS EKS 托管 Kubernetes 集群, 通过 AWS ELB 流量分发与 HPA 实现 5k RPM 下 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
- 作为 Scrum Master 主导 5 人跨职能团队, 基于 Jira 需求拆解与 Figma 高保真原型驱动开发, 需求交付准时率提高 20%

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)