

廖明月

资深机器学习工程师 · 大模型后训练 (SFT / DPO / GRPO)

5 年以上机器学习工程经验，先后任职于百度搜索与阿里巴巴大模型应用事业部，现在日本担任 LLM 技术专家。专注大模型后训练：用 GRPO 训练幻觉检测评判模型，并通过 DPO、self-reward 用评判结果提升生成模型。在医疗领域有从增量预训练、检索到生成的全链路落地经验。

工作经历

LLM 技术专家 — SERVICEMEME

2025.10 - 至今

日本

- **新闻合规筛查助手（已部署，供客户批量处理全库数据）。** 从新闻报道中自动识别贷款申请人的犯罪记录及是否为制裁对象。客户原方案为 Qwen3 + 提示词，提示词随 case 不断加长后效果明显下降，团队也没有微调经验。加入后先建立错误分类体系（标注不一致、数据未覆盖、关系挂错、关系缺失、实体边界错误）并统计各类错误的分布，再分三步改进：统一标注规范（+1.2%）、合成数据提升覆盖率（+1.5%）、以交叉方式在全量数据上构建正负样本对做 DPO，提升易错数据上的表现（+3.5%）。最终形成“扩充评测 → 构建多样性数据 → 构建正负样本 → SFT → DPO → 评测”的迭代体系。
- **面向美国医疗机构的 Text-to-Python 工具（Azure）。** 将自然语言转为数据分析 Python 代码。设计四阶段工作流：领域微调模型生成初版 → 执行并自我修正（最多 3 轮，成功率 82% → 93%）→ LLM-as-a-judge 评审 → 结合试错记录与评审结果选出最终代码。用 DPO 学习正误代码对，并扩充多表关联与低频字段数据，综合准确率提升 10%；最终代码生成准确率 90%+，分析工时减少 20% 以上。
- **奢侈品箱包 AI 鉴定系统（已上线）。** 输入一个包的多张图片，输出真伪判定和 8 个结构化属性。YOLO11s 检测 32 个鉴定点（mAP50 0.91）；真伪采用 ConvNeXtV2 / DINOv2 局部评分 + XGBoost 整包判定的三阶段方案，端到端 AUC 0.898（闭集）；官方色号用 ArcFace 度量学习做检索，新增色号无需重训。以微服务部署于 AWS，线上用于初筛，配合人工复核。

算法专家 — 阿里巴巴集团 大模型应用事业部

2023.10 - 2025.09

中国 · 杭州

产品为基于 RAG 的医学文献综述生成（两路召回 → 重排 → 大纲 → 正文），负责生成环节。

- **基于 GRPO 的幻觉检测评判模型（独立负责）。** 人工归因出前提缺失、不合理推理、张冠李戴三类错误；将长答案切分为带参考文献的 claim 逐条判断。负样本全部合成：Qwen 72B 按指定错误类型改写正确文本，并用检测模型过滤、三轮迭代得到难样本。以 GRPO 训练 Qwen 14B 评判模型（正确性奖励 + 格式奖励）。在 1000+ 条人工标注 claim 上，F1 较开源方案 RefChecker 相对提升 7%，较 SFT 基线相对提升 10% 以上。
- **RAG 生成质量优化。** 以人工评测专业性得分为指标（均为相对提升）：self-reward 对齐 +10%，claim 级幻觉率下降约 50%；用评判模型构造偏好数据做 DPO +5%；双语术语词典 +5%；RAFT 抗噪训练 +4%。
- **医疗领域增量预训练。** 挖掘医学术语，基于搜索增强合成问答与综述提纲数据，对开源基座模型做增量预训练；支撑英文文献中文摘要、文档翻译、文档问答三类线上能力。

高级算法工程师 — 北京壹永科技

2022.07 - 2023.09

中国 · 北京

- **大模型虚拟标注员（离线验证）。** 临床质控中规则无法穷举且不带语义，验证了以大模型端到端判断替代多级流水线（NER → 关系抽取 → 实体链接 → 规则）的可行性。基于 LLaMA 13B，以 QLoRA / LOMO 解决低显存微调，以病历切分 + RoPE 扩展解决 2048 上下文限制。离线准确率：疗效评价 67%→85%，手术切缘 53%→87%，放射治疗 76%→91%。
- **医疗领域增量预训练。** 实体级掩码 + SBO 目标（借鉴 SpanBERT）；以 2 层 ALBERT-tiny + span 模型替换线上 4 层 ALBERT+LSTM+CRF，推理速度提升 3 倍，模型体积缩小 30%，F1 提升 1%。
- **通用医学术语标准化。** 主导手术、诊断、检验、检查、药品 5 大值域的术语标准化，服务 1000+ 家医院；手术 / 诊断 F1 ≥ 90%，检验 / 检查准确率 ≥ 95%。

算法工程师 — 百度 搜索部2020.07 - 2021.12

中国·北京

- 搜索 Query 通用 NER。支持 70+ 实体类型、3 级层次的多标签 NER（ERNIE 预训练 + 多任务学习，高频 query 离线缓存，长尾 query 走线上 3 层 ERNIE+CRF）；小说 / 影视 / 歌曲类目较百度 UNIT 基线准确率与召回率提升 25%。
- 解析模型迁移与度晓晓。将 200+ 解析模型迁移至主搜架构；为虚拟助手度晓晓蒸馏 3 层 ERNIE 情感分类模型，线上准确率 90%+，覆盖 40% 推荐流量。

算法工程师（实习） — 贝壳找房 搜索与语言智能2019.10 - 2020.07

中国·北京

- 对话系统可控文本生成（SeqGAN + GPT）与意图分类模型上线；会话服务质量评分模型线上准确率 95.21%。

教育经历

工学硕士·软件工程 中国人民大学2017.09 - 2020.06

工学学士·物联网工程 合肥工业大学2013.09 - 2017.07

专业技能

大模型后训练：SFT、DPO、GRPO、PPO、self-reward、RAFT、增量预训练、知识蒸馏、MergeKit

RAG 与检索：幻觉检测（评判模型）、重排序、向量模型微调、文档切分策略

框架与工具：PyTorch、TensorFlow、PaddlePaddle、Transformers、vLLM、Python、C++、Java

云与基础设施：GCP（Cloud Run、Vertex AI、IAM、Terraform）、Azure AI Foundry、Salesforce API

模型：Qwen 系列（含 RTX 5090 本地部署）、LLaMA、ERNIE、BERT / ALBERT 系列、GPT-4（作为 teacher / judge）

获奖情况

- 全国研究生数学建模竞赛 **三等奖**（2018）
- 神策杯关键词预测大赛 **初赛 5/641**（2018）