

# 求职简历

PERSONAL RESUME

我一直在努力！

## 基本信息

姓名： 闫树妹  
年龄： 32  
学历： 大学本科  
求职意向： nlp 工程师

手机： 15102131535  
邮箱： 1075827662@qq.com  
微信： yanshumei124  
地址： 上海市浦东新区



## 专业技能

掌握 Bert、baichuan、chatGLM 的 lora, sft, p-tuning 等微调技术；  
掌握 PyTorch 框架，熟悉深度学习网络结构 transformer，熟悉 langchain, langgraph；  
掌握 NLP 中数据处理、数据清洗等；  
掌握提示词工程的几种模式，ai agent  
掌握 Python 以及 Numpy、Pandas、SciPy 等科学计算库框架；  
熟悉 Linux 系统、Docker、SqlServer、GP、mysql 数据库命令；了解 jenkins, k8s 部署

## 教育背景

海洋生物学	广东海洋大学（统招一本）	2010 年-2014 年
金融学	上海财经大学（在职硕士 211）	2024 年-至今

## 工作经历

研发经理（nlp）	上海柯林布瑞信息技术有限公司	2018.12-至今
自然语言处理工程师	上海安捷致善医学科技有限公司	2017-2018
生物信息工程师	苏州贝斯派信息技术有限公司	2015-2017

## 工作经验

项目名称：基于 LLM 的医疗垂直领域模型微调（**华西医院检索智能体**）  
**技术栈**：PyTorch/Transformers/Deepspeed/Peft/LoRA/Docker/k8s/langGraph/  
**模型基座**：Baichuan/ChatGLM  
**周期**：2023.5-至今

- 项目背景  
满足医生护士科研一条龙服务  
解决通用大模型在医疗检索中存在的幻觉问题及本地化部署。  
医生科研对话流程智能体
- 核心工作  
**数据收集**：导出 solr 中的检索记录+权威医学文献+临床指南等。  
**清洗增强**：去重（相似度过滤检索语句中相似的语句）、标准化（术语统一）。  
构建 Prompt 模板  
**参数高效微调（PEFT）**：采用 LoRA（低秩适配），仅训练 0.1% 参数量。  
**对比调研**：LoRA vs. Adapter，最终选择 LoRA（实现简单、接近全参数调参）。

**超参调优：**学习率：通过 LR Finder 确定三角循环学习率（峰值  $3e-5$ ）。

Batch Size：动态调整（8→32，依显存占用）。

过拟合：Early Stopping。

**检索增强生成：**引入了院内资产表，使用 RAG 技术，快速检索，减少大模型幻觉问题

**评估与迭代量化指标：**专业术语准确率（人工评估+BLEU-4）。

幻觉率（基于 NLI 模型检测矛盾大于 0.8）。

**多智能体协作架构：**使用 LangGraph 实现搜索流程的有限状态机控制以及多轮对话。

**动态路由设计：**根据查询意图（辅助诊断/检索患者/检索文献/检索指南）自动选择特定 Agent 以及对应的检索库。

**Prompt 模板：**分别为几种意图设置 prompt 模板，方便医护人员检索

**离线部署：**打包镜像离线部署到院内（或者项目经理通过应用商店下载，dockerfile 已上传到 jenkins，k8s 部署适配）

**模型上线后，好评率提升 95%。**

项目经验二：单病种 ai agent

### 1、项目背景

实现项目现场的结构化任务，完成国家对电子病历互连互通和病种库上报的要求

### 2、核心工作

首先分析不同文本针对不同文本设计满足各个产品的结构化要求的 prompt 语句，设计 langchain 工作流，然后对接 etl

项目经验三：BI 运营大模型平台

项目描述：

为帮助医院管理者及时了解院内的运营现状，便于科室主任实时了解科室相关指标，基于大模型和 DataInsight 开发的 BI 平台，功能包含多轮对话、根因分析、数据分析、数据预测等。

个人职责：

1、基于用户之前的常用点击操作，及 DataInsight 中各个指标、维度信息，准备数据集；

2、基于多轮问答，完成用户的下钻操作；

3、大模型实现数据分析、数据预测等功能，以及将提取的指标、起始和结束时间、图形、科室、年、月等维度信息输入给 DataInsight，进行环比、同比、各个图形等的展示。

以上是我的三个经典项目，敬请查阅，别的项目还很多，核心内容由以上总结