**企业实习工作日志**

|  |  |
| --- | --- |
| 学生学院： | 重庆大学计算机科学与技术学院 |
| 专业班级： |  |
| 学生学号： |  |
| 学生姓名： |  |

**四川华迪信息技术有限公司制**

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2025年 06月 21日 | 实习地点 | DS3401 |
| 实习内容 | 1. 完成训前测试 2. 学习Numpy的综合使用。 3. 确定选题 4. 学习相关知识 5. 收集相关数据 | | |
| 工作开展情况 | **6.21上午**  **完成老师给定的训前测试。**  **学习Numpy的综合使用。**  **6.21下午**  **确定选题：大模型智能辅诊系统** 内容：   1. 数据采集 数据来源：医院HIS系统、PACS影像系统、可穿戴设备实时数据、医学文献库。 2. 数据类型：结构化数据（检验指标、用药记录）、非结构化文本（门诊病历、影像报告）、时序数据（心电监护、血糖监测）、高维数据（CT/MRI DICOM文件） 3. 数据处理：多模态对齐：医学文本去标识化（HIPAA合规）、DICOM图像窗宽窗位标准化、穿戴设备数据异常值检测特征工程：使用BioBERT提取临床文本特征、ResNet-50提取影像特征、时间序列特征提取（STL分解） 4. 知识注入：UMLS医学本体实体链接、临床路径规则嵌入模型 5. 设计微调方案：LLaMA-Med在百万级电子病历上的继续预训练、多任务学习框架（诊断预测+报告生成）、检索增强生成（RAG）结合UpToDate知识库 6. 推理优化：思维链（CoT）提示工程、不确定性量化（蒙特卡洛Dropout）、可解释性模块（显著性热力图） 7. 应用模块辅助诊断：生成Top-5鉴别诊断及置信度、自动关联检验检查建议报告生成：结构化影像报告自动生成、门诊病历语义完整性检查 8. 病程管理：自动提取SOAP病历要素、生成个性化随访方案   **学习相关知识，收集相关数据，确定方向** 用毛选微调大模型的全流程  https://www.bilibili.com/video/BV1Gn7nz3EnG/?share\_source=copy\_web&vd\_source=a991aeb9c9e2813baf902a4165833bb9  文本To问题 智能分割转换工具  https://github.com/ConardLi/easy-dataset/blob/main/README.zh-CN.md  Qwen3 企业级微调实战【4小时】  https://www.bilibili.com/video/BV1YLE1zyEvX?spm\_id\_from=333.788.videopod.episodes&vd\_source=14149f51bbbd96bd5d881a30ef958585  DeepSeek-R1医学领域微调实战  https://zhuanlan.zhihu.com/p/31610343745  本地部署后在线/后端调用大模型  https://www.cnblogs.com/sio2zyh/p/18930757 | | |
| 工作反思与总结 | 今天通过训前测试与Numpy学习，掌握了基础的数据处理能力。确定了“大模型智能辅诊系统”的选题后，结合医学领域数据特点，梳理了完整的项目结构与实现路径，对后续的开发工作提供了清晰指引。通过资料收集与系统设计，初步建立了多模态输入与模型微调的认知框架。 | | |
| 下一步工作计划 | 1. 学习如何微调 2. 收集微调相关信息 3. 前后端选型 4. 确定项目需求 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 6 月 22日 | 实习地点 | DS3401 |
| 实习内容 | 1. 学习如何微调 2. 收集微调相关信息 3. 前后端选型 4. 确定项目需求 | | |
| 工作开展情况 | **6.22上午**  **学习如何微调**   1. 收集数据，将其进行分割并总结为问答形式 2. 将问答形式以Json格式导出 3. 以Json格式的问答对模型进行微调 4. 微调后验证微调效果 5. 参考视频：https://www.bilibili.com/video/BV1Gn7nz3EnG/?share\_source=copy\_web&vd\_source=a991aeb9c9e2813baf902a4165833bb9   **收集微调相关信息**  **DeepSeek-R1医学领域微调实战**  https://zhuanlan.zhihu.com/p/31610343745  **本地部署后在线/后端调用大模型**  https://www.cnblogs.com/sio2zyh/p/18930757  **6.22下午**  **前后端选型**  前端：Streamlit/Gradio  可以参考赛博华佗：https://github.com/Warma10032/cyber-doctor  前端使用的技术即为Gradio  后端：Python（FastAPI）  **确定项目需求**  我们要做的是**大模型智能辅诊系统**  视角应该是：系统使用者，也就是“我”，是一名医生，我通过与病人交流，将其病历以及描述（比如说我头疼喉咙痛），交给AI，让AI来判断病人得了什么病，并为其推荐几种药物出来。  我通过AI给出的建议，进行病情和药物的选取，模拟在医院中，医生听取病人意见后在系统上进行病情诊断和药物选取的过程，这个病情诊断和药物选取的过程交由AI处理，我只需要在AI给出建议后（比如3种病和10种药中）进行选择，例如选择感冒——布洛芬这种组合，然后点击按钮生成诊断报告  整体流程如图：  从技术上来说，整体流程应该是： | | |
| 工作反思与总结 | 今天重点学习了大模型微调的基本流程，并尝试将问答数据整理为标准格式，为后续训练打下基础。同时完成前后端框架选型，为系统部署方向提供依据。对项目场景有了更深入理解，也逐步明确了核心模块的职责划分。 | | |
| 下一步工作计划 | 1. 配置环境 2. 测试模型效果 3. 寻找数据集 4. 对模型进行初步微调 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 6 月 23日 | 实习地点 | DS3401 |
| 实习内容 | 1. 配置环境 2. 测试模型效果 3. 寻找数据集 4. 对模型进行初步微调 | | |
| 工作开展情况 | **6.23上午**  **配置环境**   1. 安装CUDA      1. 安装CUDA Toolkit      1. 安装PyTorch   pip3 install torch torchvision torchaudio --index-url [https://download.pytorch.org/whl/cu](https://download.pytorch.org/whl/cu126)128   1. 下载Qwen3 1.7B 量化模型进行先行测试   https://www.modelscope.cn/models/unsloth/Qwen3-1.7B-unsloth-bnb-4bit  **测试Qwen3 1.7B小模型**  修改system\_prompt对Qwen3模型进行修改      **6.23下午**  **寻找小规模法律数据集对Qwen3进行初步微调**      处理成适合微调的格式    进行小规模训练（训练批次很小，为了验证微调步骤准确性）    对微调后的模型进行验证      发现已经可以回答数据集中的问答了，但是其他方面的性能收到了影响。 | | |
| 工作反思与总结 | 完成了环境配置与模型部署的关键步骤，顺利跑通Qwen3模型测试。通过小规模数据集对模型进行初步微调，掌握了基本训练流程。虽然存在性能损耗问题，但为后续的微调与验证积累了实践经验。 | | |
| 下一步工作计划 | 1. 对模型重新微调 2. 保存模型 3. 验证模型效果 4. 选择更大的数据集 5. 重新微调 6. 学习Gradio 7. 制作GradioDemo | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 6 月 24日 | 实习地点 | DS3401 |
| 实习内容 | 1. 对模型重新微调 2. 保存模型 3. 验证模型效果 4. 选择更大的数据集 5. 重新微调 6. 学习Gradio 7. 制作GradioDemo | | |
| 工作开展情况 | **6.24上午**  **重新微调**  昨天微调是在0.6B+100条数据的数据集下训练的，batch\_size也只有1，目的只是为了熟悉微调流程。 在过完整个流程后，今天上午在Qwen-4B模型下对一万条数据的数据集进行微调，微调效果很好，模型在保留原本知识的情况下依旧能够对数据集中的问答进行正确的回答。  **保存模型**  微调完数据集后，需要对其进行保存，毕竟不能每次使用之前都进行微调，因此需要探索能够保持高精度保存的方法。  经过研究发现，进行高精度保存会出现无法解决的错误，在Github上进行追踪调查应该是某次PR对保存方法进行了修改，导致出现了这个问题。在issues中进行浏览发现都是近两周才出现的问题。  问题：input tensor size [6291456, 1] and output tensor size [2048, 6144] should match. 相关issue链接：https://github.com/unslothai/unsloth/issues/2749  **至今无法解决。**  因此只能退而求其次，选择4bit精度保存。 参考文章：https://blog.csdn.net/BIT\_666/article/details/132065177  先使用model = model.merge\_and\_unload()，对模型进行合并和去lora层，这个方法有UserWarning: Merge lora module to 4-bit linear may get different generations due to rounding errors. 也就是刚刚提到的低精度保存，虽然会缺失精度，但暂时没有办法。  之后进行常规方法保存   |  | | --- | | Python # 对去掉Lora层的model进行保存 model.save\_pretrained("Qwen3-4B-finetuned-fp16") tokenizer.save\_pretrained("Qwen3-4B-finetuned-fp16") |     保存后对模型大小进行查看。      发现大小没有明显变化，即合并保存完毕。  **6.24下午**  **验证保存的模型是否保存成功**  虽然保存成功了，但不知道是否有具体的变化，即不知道再次部署模型能否按照数据集中的内容进行回答，因此需要进行测试  部署后模型失去了微调后的能力，应该是中间的某些步骤出了问题，因此需要重新选择数据集进行测试。  **在更大的医学数据集上重新微调**  为了验证是否是微调过程出现问题，这里选择一个更大的医学数据集进行验证。  **学习Gradio，尝试用已有的模型做一个在线翻译器**  用Gradio做一个在线页面，调用本地部署的模型，做一个在线翻译器。  用Gradio提供的Demo和部分代码，结合本地部署的Qwen3-4B模型，能够实现输入英文自动翻译中文。 | | |
| 工作反思与总结 | 今天在更大数据集上完成了正式微调，并深入研究了模型保存问题。尽管高精度保存仍存在技术障碍，但已通过4bit精度成功保存模型。利用Gradio构建了简单Demo，验证了模型能力与前端交互效果，为进一步系统搭建奠定基础。 | | |
| 下一步工作计划 | 1. 学习Transformer原理 2. 更换数据集继续微调 3. 继续学习Gradio 4. 学习RAG相关内容 5. 学习Gradio相关排版内容 6. 收集医学大模型评测体系相关内容 7. 学习如何使用病历模板生成病历 8. 学习如何将Word病历转为PDF 9. 利用Gradio在线生成病历PDF并提供下载 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 6 月 25日 | 实习地点 | DS3401 |
| 实习内容 | 1. 学习Transformer原理 2. 更换数据集继续微调 3. 继续学习Gradio 4. 学习RAG相关内容 5. 学习Gradio相关排版内容 6. 收集医学大模型评测体系相关内容 7. 学习如何使用病历模板生成病历 8. 学习如何将Word病历转为PDF 9. 利用Gradio在线生成病历PDF并提供下载 | | |
| 工作开展情况 | **6.25上午**  **学习Transformer原理**  为更好地理解人工智能的原理，学习Transformer和注意力机制等。 https://www.bilibili.com/video/BV1XH4y1T76e/?spm\_id\_from=333.337.search-card.all.click&vd\_source=14149f51bbbd96bd5d881a30ef958585  **更换数据集继续微调模型**  由于昨天微调保存后，模型失去了微调给予的新能力，因此要在更大的数据集上继续微调。  微调后发现效果很好，能够对数据集中提出的问题给出准确地回答      如果是Qwen-4B原本的模型则不能正确回答  **继续学习Gradio**  能够在线访问本地部署的Qwen3-4B，并将微调后的模型与其对比，验证微调效果。  **6.25下午**  **学习RAG相关内容**  https://www.bilibili.com/video/BV19RJhzyEWN/  **学习Gradio相关排版功能**  为了做出更精美的前端页面，学习Gradio相关排版功能，学习如何编写一个ChatBot页面。 https://www.gradio.app/docs/gradio/chatbot  **收集医学大模型评测体系相关内容**  为了验证微调后大模型对于医学内容回答的效果，收集相关测评体系内容，方便后续进行测试。  **学习如何使用病历模板生成病历并转为PDF**   1. Word模板填充：使用python-docx-template库 2. 使用win32com（依赖Microsoft Word）         能够根据给定内容，填写模板并生成Word和PDF  **利用Gradio在线生成病历PDF并提供下载**  使用Gradio做了一个Demo，能够输入相关信息，自动填入模板并生成PDF提供下载。 | | |
| 工作反思与总结 | 今天继续推进了微调实验并取得理想结果，模型对定制数据集反应良好。同时学习Transformer原理、Gradio排版与RAG机制，加深了对大模型本质与问答系统架构的理解。完成病历模板填充与PDF生成功能的原型，功能实现进展顺利。 | | |
| 下一步工作计划 | 1. 在Gradio页面中接入大模型 2. 尝试在Gradio中使用API接口 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 6 月 26日 | 实习地点 | DS3401 |
| 实习内容 | 1. 在Gradio页面中接入大模型 2. 尝试在Gradio中使用API接口 | | |
| 工作开展情况 | **6.26上午**  **在Gradio页面中接入大模型**  尝试接入Qwen3-4B原始模型测试效果，看能否构建一个对话页面，可以根据大模型输出生成病历    目前已成功接入，支持根据问题进行指定格式的回答。  填入信息后能够正常生成Word和PDF格式的病历。    但目前Word格式存在部分问题，需要进行调整。  调整完毕后可以正常对话且生成对应病历  **6.26下午**  **尝试在Gradio中使用API接口**  在本地运行大模型的情况下，使用Gradio可以正常对话，但每次调用模型时都需要重新启动模型，速度很慢。并且最后进行部署时使用的也是本地部署模型的API，因此需要尝试使用API调用模型，提前适配。  这里使用百炼提供的Qwen3-API，尝试接入API后是否可用，为本地API做铺垫。  后发现使用本地Ollama进行模型部署时，会自动生成一个localhost的API，可以直接request请求使用。  因而没有使用QwenAPI进行尝试，直接转用本地API，发现效果良好 | | |
| 工作反思与总结 | 今天将本地模型成功嵌入Gradio对话页面，完成病历生成的初步闭环。尝试本地API调用以替代模型反复加载，极大提升了响应速度，为部署提供技术储备。系统交互与页面功能进一步成熟。 | | |
| 下一步工作计划 | 1. 继续优化Gradio页面 2. 将Gradio项目同步至Github 3. 添加语音转文字功能 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 6 月 27日 | 实习地点 | DS3401 |
| 实习内容 | 1. 继续优化Gradio页面 2. 将Gradio项目同步至Github 3. 添加语音转文字功能 | | |
| 工作开展情况 | **6.27上午**  **继续优化Gradio页面**  继续优化Gradio页面，使其在本地接口支持下，添加更多功能。  目前的想法是，实现一个功能：AI生成之后，点击一下AI的对话记录，将其记录到备选框中，可以手动修改，修改过后点击确认生成PDF。  PDF功能已实现，能够根据大模型回答生成可编辑的文字。    右侧部分可以进行修改，修改完毕后会根据修改内容生成病历PDF。  **将Gradio迁移至Github**  将Gradio页面迁移至Github，方便团队协作与进度同步  **6.27下午**  **添加语音转文字功能**  尝试添加语音转文字功能，可以支持语音输入内容。    一开始使用的是Whisper，openai的开源模型，对于中文识别效果很差，后续需要更换适配中文的模型。 （这里说的原句是：我头痛，我怎么了），Whisper识别正确率几乎为0  明天尝试换一个模型 | | |
| 工作反思与总结 | 今天在Gradio页面中实现了可编辑输出与PDF保存功能，提升了用户交互体验。完成登录功能原型设计，并尝试将项目同步至Github，为协作开发提供便利。语音输入初步接入，但识别效果仍需优化。 | | |
| 下一步工作计划 | 1. 更换语音转文字模型 2. 修改页面布局 3. 添加图片上传 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 6 月 28日 | 实习地点 | DS3401 |
| 实习内容 | 1. 更换语音转文字模型 2. 修改页面布局 3. 添加图片上传 | | |
| 工作开展情况 | **6.28上午**  **更换语音转文字模型**  昨天语音转文字的功能特别差，今天要换一个新模型  目前准备使用阿里通义的SenseVoice    更换模型之后检验效果    发现识别效果较好，识别正确率很高。  **6.28下午**  **修改页面布局**    对页面布局进行了重构，看起来更加清爽简洁  **添加图片上传**    增加了一个Tab栏，可以支持在文本诊疗和图像处理之间进行转换。  方便后续影像识别使用。 | | |
| 工作反思与总结 | 今天成功更换了更适配中文识别的语音模型，识别准确率显著提升。对页面布局进行优化，引入影像识别模块并支持图片上传，为后续医学影像分析功能预留接口，系统整体功能性与可扩展性持续增强。 | | |
| 下一步工作计划 | 1. 尝试添加登录页 2. 优化代码 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 6 月 29日 | 实习地点 | DS3401 |
| 实习内容 | 1. 尝试添加登录页 2. 优化代码 | | |
| 工作开展情况 | **6.29上午**  **尝试添加登录页**  需要添加用户登录的功能，因此需要接入数据库，暂时决定选择SQLite，比较轻量且好用  用户登录后，可以查看之前生成过的病历。  为了不破坏原来代码的功能性，先写一个另外的Demo测试效果（页面布局和视觉效果先不做）    支持登录和注册，注册的用户数据会存入SQLite数据库    登入系统后会显示主界面    支持显示历史聊天和查询曾经生成过的文件（对应到实际的项目中，应该是以前生成过的病历）  生成的病历文件格式为PDF，在数据库中需要用二进制保存。    可以支持下载已有的文件  **6.29下午**  **优化代码**  继续对已有代码进行优化，对冗余代码进行删除，保留关键部分。  并与原本的Gradio代码进行整合，使其支持原有功能。  完成了生成病历保存至数据库，可以按照用户查询历史病历。 | | |
| 工作反思与总结 | 今天完成登录与注册功能并接入SQLite，用户管理功能初步成型。优化代码结构与界面布局，支持生成病历保存与历史病历查询下载，系统功能开始向真实应用靠拢，用户体验进一步提升。 | | |
| 下一步工作计划 | 1. 提供下载病历功能 2. 完成Gradio页面需求 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 6 月 30日 | 实习地点 | DS3401 |
| 实习内容 | 1. 提供下载病历功能 2. 完成Gradio页面需求 | | |
| 工作开展情况 | **6.30上午**  **提供下载**  昨天只完成了生成病历保存和病历查询，还差一个历史病历下载的功能没有完成，今天继续完成。    完成后，可以点击【下载】，下载历史病历。  **6.30下午**  **完成Gradio页面需求**   1. 重构项目结构     把操作逻辑相关的函数都剥离了GradioDemo   1. 修改登录页面，登录页和注册页分开 2. 添加退出登录按钮 3. 主治医师换成用户名     添加了对主治医师的修改支持（即username） | | |
| 工作反思与总结 | 今天完善了历史病历的下载功能，重构项目结构与登录逻辑，细化权限控制与界面显示。使系统内容更加规范化与专业化。 | | |
| 下一步工作计划 | 1. 对WebUI进行修改 2. 修改登录页 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 7 月 2日 | 实习地点 | DS3401 |
| 实习内容 | 1. 对WebUI进行修改 2. 修改登录页 | | |
| 工作开展情况 | **7.2上午**  6.30下午的需求还有两个没做，分别是：   1. 修改登录页面，登录页和注册页分开 2. 添加退出登录按钮   **WebUI修改-0702版**  有了新的需求：   1. 【部分完成】登录注册页面样式和使用逻辑修改（按照之前所说的样式，背景可以选一张精致大图）      1. 【已完成】© 2025 智能医疗诊断系统 | 版权所有 改为 © 2025 智渝——重大智慧医疗辅诊系统 | 版权所有      1. 【已完成】中间大标题改为：智渝——重大智慧医疗辅诊系统；最右侧加上“退出登录”功能选项      1. 【已完成】这个语音识别好像无法识别与添加标点？好像目前的语音识别可以实现，这里尽量修改优化一下。      1. 【已完成】生成PDF的进度条可以加上，不然等待时候没有反馈，不直观。      1. 根据项目要求，病例输出的格式要大改，具体见下：      |  | | --- | | 其实我觉得我们现在写的已经差不多符合了，但是表述上要贴合课程要求的。 |   病历要符合SOAP格式，分为：  S (Subjective - 主观信息)：从医患对话记录中（模型输入）提取如主诉、现病史、既往史、个人史、家族史、过敏史等关键主观信息。  O (Objective - 客观信息)：整合患者所有可测量的、医生观察或仪器检测的客观数据，比如体温、血压、心率等客观生命体征或者血常规、生化、凝血功能、病原学检测等实验室检查结果或者其它。  A (Assessment - 评估)：要**给出Top-5 鉴别诊断**，基于整合的S和O信息生成最可能的5个诊断（或问题列表），为每个诊断量化或定性（高/中/低）地给出置信度（概率或置信分数），对于每个置信度的提出，要附带给出简短的推理依据。  P (Plan - 计划)：一方面是**给出个性化随访方案**，基于之前的患者具体情况（S和O的信息），与当前诊断（A的信息），动态生成个体化的随访建议，内容包括如随访时间点（如：3天后复诊、1周后复查血常规、1月后复查影像）、随访目的（评估疗效、监测副作用、调整方案）、随访内容（需要复查哪些项目、关注哪些症状变化）、患者教育要点（如：出现XXX症状需立即就医）。另一方面**给出自动关联检验检查建议**，为支持或排除Top鉴别诊断，给出建议下一步需要进行的关键实验室检查（血常规、生化、特定标志物等）、影像学检查（X光、CT、MRI、超声等）或者其他检查（病理活检、内镜、功能检查等）。最好指明每项检查是为了“确诊XX”、“排除XX”、“监测XX副作用”。  最后基于以上所有，可给医生提供初步的药物或非药物治疗方案建议，如病情复杂或超出当前科室范围，也可建议请相关科室会诊或转诊。   1. 【已完成】最后PDF的内容根据6中的信息对应修改。病历标题改为“重庆大学附属计科医院”，主治医师可绑定“账号名”，就诊时间绑定当日时间，下面的内容依照SOAP格式呈现。     **7.2下午**  **继续修改登录页**   1. 给登录页加一个背景大图      1. 修改了图标颜色，防止被背景影响 2. 处理一下门诊病历的模板 3. 修改了prompt，让模型输出更贴合SOAP要求 | | |
| 工作反思与总结 | 今天根据反馈优化登录界面与交互细节，页面美观性与功能性并进。调整大模型prompt与病历显示逻辑，使生成内容更贴合医学标准，也方便医生查看与编辑，提高系统实用价值。 | | |
| 下一步工作计划 | 1. 修改prompt 2. 调整显示部分 3. 医学影像报告生成 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 7 月 3日 | 实习地点 | DS3401 |
| 实习内容 | 1. 修改prompt 2. 调整显示部分 3. 医学影像报告生成 | | |
| 工作开展情况 | **7.3上午**  **修改prompt**    使其更符合项目要求，以SOAP格式输出。  **调整显示部分**    大模型给出的回复并不全部显示在病历中，只取关键部分显示    全部的回答会显示在最左边    右边的病历编辑部分只会显示最重要的信息，方便生成病历PDF    可完整生成病历  **7.3下午**  **医学影像报告生成**  修改Gradio UI界面，使其适配医学影像生成 | | |
| 工作反思与总结 | 今天优化了prompt设计与内容截取逻辑，提升生成病历的规范性与准确性。调整大模型prompt与病历显示逻辑，使生成内容更贴合医学标准，也方便医生查看与编辑，提高系统实用价值。 | | |
| 下一步工作计划 | 1. 微调UI界面 2. 完善医学影像页面 3. 设计医学影像模板 4. 完善README 5. 接入医学影像模型 6. 撰写新的提示词 7. 重构医学影像页面 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 7 月 4日 | 实习地点 | DS3401 |
| 实习内容 | 1. 微调UI界面 2. 完善医学影像页面 3. 设计医学影像模板 4. 完善README 5. 接入医学影像模型 6. 撰写新的提示词 7. 重构医学影像页面 | | |
| 工作开展情况 | **7.4上午**  **微调UI界面**  继续微调UI界面，使其风格统一且样式美观    **医学影像页面**    暂时为医学影像页面添加UI，后续会继续优化  **医学影像模板**  为医学影像报告设计了一个模板    **完善了README**  添加了一些信息，方便后续开发者构建项目  **7.4下午**  **医学影像模型接入**  接入医学影像模型    **提示词撰写**  为新接入的医学影像模型撰写提示词，使其能够正确分析影像并按格式输出    **重构医学影像页面**  在Gradio中重构医学影像页面 | | |
| 工作反思与总结 | 今天完成了医学影像页面UI设计与影像模板构建，并成功接入模型及定制提示词。通过持续UI与提示词调优，初步形成医学图文一体的诊疗能力，系统整体完成度再进一步。 | | |
| 下一步工作计划 | 1. 继续完善医学影像页面 2. 检查生成的医学影像报告是否正常 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 7 月 5日 | 实习地点 | DS3401 |
| 实习内容 | 1. 继续完善医学影像页面 2. 检查生成的医学影像报告是否正常 3. 添加知识库页面 4. 修正部分文本 5. 修改CSS页面 | | |
| 工作开展情况 | **7.5上午**  **医学影像页面**  对昨天未能完成的医学影像页面继续做修改    **医学影像报告**  查看生成的医学影像报告    **7.5下午**  **添加知识库页面**    支持上传文件到知识库，查看知识库中已有的文件。  **修正部分文本**  对错误或表述不恰当的文本进行修正  **修改CSS页面**  调整了部分文字的显示样式 | | |
| 工作反思与总结 | 医学影像功能初步可用，完成影像报告生成逻辑，添加知识库页面支持上传与预览。进一步修正展示细节与样式，整体项目趋于完善，体验感持续优化。 | | |
| 下一步工作计划 | 1. 验证医学影像模型的可用性 2. 修改Prompt 3. 验证大模型效果 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 7 月 6日 | 实习地点 | DS3401 |
| 实习内容 | 1. 验证医学影像模型的可用性 2. 修改Prompt 3. 验证大模型效果 | | |
| 工作开展情况 | **7.6上午**  **验证医学影像模型的可用性**  接入医学影像模型，进行判断和验证，查看生成报告的效果  验证后发现生成的报告的代码有一定的问题，对问题进行了修复，目前可以正常生成  **7.6下午**  **修改Prompt**  医学报告大模型依然有一定问题，不能按照既定格式输出，导致截取内容很难，因此需要对大模型进行一定调整。  **验证大模型效果**  对调整后的模型进行测试，幻觉少了很多，大多数情况下可以标准输出 | | |
| 工作反思与总结 | 调整影像模型Prompt，改善模型输出格式问题，确保信息结构清晰、准确。大模型在医学问答方面的幻觉问题显著减少，核心问答能力初步稳定。 | | |
| 下一步工作计划 | 1. 对错误文本进行修正 2. 添加提示信息 3. 接入MCP | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 7 月 7日 | 实习地点 | DS3401 |
| 实习内容 | 1. 对错误文本进行修正 2. 添加提示信息 3. 接入MCP | | |
| 工作开展情况 | **7.7上午**  **对部分文本进行修正**  修正了部分错误文本  **添加提示信息**  为部分内容添加提示信息        **7.7下午**  **接入MCP**  支持深度思考和联网搜索，联网时会根据联网内容进行一定调整来回答        客观内容中，已经显现了联网搜索的内容。 | | |
| 工作反思与总结 | 接入MCP模块，模型具备联网搜索与深度思考能力，显著提升回答准确性与客观性。同时在界面中引入提示机制，增强用户使用体验，系统智能性进一步增强。 | | |
| 下一步工作计划 | 1. 接入RAG 2. 加入知识库内容预览功能 3. 测试 4. 修复BUG | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 7 月 8日 | 实习地点 | DS3401 |
| 实习内容 | 1. 接入RAG 2. 加入知识库内容预览功能 3. 测试 4. 修复BUG | | |
| 工作开展情况 | **接入RAG**  加入检索增强功能，可以根据自建知识库进行问答      同时API也已调通，只需要接入即可。  **加入知识库内容预览功能**  在实际模型使用之前，可以先根据知识库已上传文件的内容对模型进行调用，查看使用效果，使用满意了再启用检索增强。    可以输入想要检索的内容，查看检索增强的效果，满意后再使用。    预览效果不错，用到了知识库中的内容。  **7.8下午**  **测试**  对项目进行结尾，对大部分功能进行测试，后续会开展更为详细的测试。  **修复bug**  对测试过程中发现的bug进行修复。 | | |
| 工作反思与总结 | 成功接入RAG模块，实现基于自建知识库的智能问答功能。新增知识库内容预览功能，提升模型调用前的可控性和可解释性。整体系统形成“知识输入—推理生成—PDF输出”的闭环流程。 | | |
| 下一步工作计划 | 1. 继续测试 2. 继续修复BUG 3. 文书工作 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 7 月 9日 | 实习地点 | DS3401 |
| 实习内容 | 1. 继续测试 2. 继续修复BUG 3. 文书工作 | | |
| 工作开展情况 | **7.9上午**  **测试**  继续对项目开展测试，发现了一些小问题  **修复bug**  对小问题进行修复    **文书工作**  写项目相关总结，完善文档  **7.9下午**  **文书工作**  继续完成需要写的文档。 | | |
| 工作反思与总结 | 对项目进行整体测试与Bug修复，保障系统稳定性与完整性。同时开展文书整理与总结，为项目验收与后期迭代做好准备，整个开发流程趋于收尾阶段。 | | |
| 下一步工作计划 | 1. 录制视频 2. 讨论细节 3. 准备答辩 | | |

**工作日志**

|  |  |  |  |
| --- | --- | --- | --- |
| 工作日期 | 2024 年 7 月 10日 | 实习地点 | DS3401 |
| 实习内容 | 1. 录制视频 2. 讨论细节 3. 准备答辩 | | |
| 工作开展情况 | 对录制的演示视频进行审查。  对所有事情进行最后的收尾。  完善项目细节。  准备一下明天答辩可能会用到的东西。 | | |
| 工作反思与总结 | 本次为期三周的实训经历，是我本科阶段理论与实践结合最紧密、技术挑战最丰富的一次综合训练。整个过程中，我不仅从0构建起一个完整可用的智慧医疗辅助诊断系统，还深入参与到系统各核心模块的设计、实现与测试环节，全面提升了个人的工程能力、问题解决能力和项目管理意识，收获颇丰。 | | |
| 下一步工作计划 | 实习结束，无进一步计划。 | | |