

**企 业 项 目 实 践**

**课 程 任 务 书**

题目： 基于大模型技术的变压器实验判据生成

姓 名： 武炫晔

学 号： U202210866

同组成员： 李权韬、林皓闻

班 级：机械本硕博 2022 班

（任务起止日期： 2025 年 3 月 1 日 ～ 2025 年 6 月 30 日）

1. 题目来源

请填写企业信息及题目来源于哪个工程项目或者企业需求或者社会需求。

国网电力科学研究院武汉南瑞有限责任公司

“立足链式思维探索服务合同管控新业态研究与应用”

1. 实践目标

请填写具体的实践目标

本课题面向电网关键设备-变压器入网前实验检测要求，实现相应国标实验判据自动提取功能，并开发软件原型系统，具体包括国标文档读取的大模型微调、设备实验判据自动化提取、实验判据的知识图谱构建、交互操作的UI设计和结果可视化显示。

1. 实践内容

请填写具体的实践内容

1. 选定合适的开源OCR工具，并针对试验标准文件中的深色背景、干扰元素等进行适当处理，以确保OCR的识别效果达到最佳；
2. 实现试验标准文件中的文字、图片、表格、CAD示意图和公式等多模态信息的抽取，以及关联和对齐，以支持后续知识图谱的构建；
3. 完成Deepseek等开源大模型的本地化部署，搭建大模型框架环境。
4. 任务与分工

武炫晔(组长)：负责标准文件读取识别和大模型框架部署，以及整体项目推进；

李权韬：负责构建知识图谱和RAG；

林皓闻：负责UI界面设计开发和问答助手。

五、已具备的实践条件

1、 请填写企业条件和课题组能提供的条件。

1）为参加课题的学生提供数据支持，以及高性能计算服务器资源和技术指导；

2）参与课题的学生需要在课题结束答辩时，提交整个课题学习周期的答辩报告归档所属学校，包括但不限于由此过程积累而产生的研制报告、专利及著作权等知识产权、教材支撑资料等。

六、进程安排

|  |  |  |  |
| --- | --- | --- | --- |
| 设计阶段 | 设计内容摘要 | 周数 | 备 注 |
| 选题 | 基于大模型技术的变压器实验判据生成 | 第2周 |  |
| 详细方案设计 | OCR工具和大模型选型及框架设计 | 第3周 |  |
| 具体企业项目训练 | 1）技术选型与学习准备  2）标准文件的多模态信息自动化抽取 | 第4-13周 |  |
| 总结和文档整理 | 1. 汇总整理项目过程文档  2. 总结团队合作和项目管理经验；  3. 撰写项目报告书和个人总结。 | 第14-15 周 |  |
| 答辩 | 答辩展示及导师评分。 | 第 16周 |  |

七、实践成果要求

1. 实践总结报告 1本；

2. 项目要求的成果，可以是大创项目，专利，基金，程序软件，实验结果，实物作品，论文等；

八、参考资料

[1] 知识图谱与 RAG 模型结合学习路线（已发给学生）

指导教师： 胡中旭

企业导师： 李林

2025年 2月 25 日