**南方科技大学本科生毕业设计（论文）任务书**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 设计（论文）题目 | 基于大语言模型的毕业设计评分系统研发 | | | | | | |
| 学生姓名 | 王谦益 | 学号 | 12111003 | 专业 | 计算机科学与技术 | 系/研究中心 | 计算机科学与工程系 |
| 主要任务及基本要求（包括设计或研究的内容、要求与指标、应完成的成果、进程安排及主要参考文献目录等）：  在研究前期会优先处理数据集，通过自动化或手动筛选步骤，过滤出完整的PDF报告，去除缺失评分或缺少内容的文件，确保数据集的完整性和一致性。对筛选出的PDF报告进行文字提取，以便后续的文本分析和评分。采用OCR或其他文本提取工具，确保获取高质量的文本数据。将提取的文字格式化，去除页面布局和格式的影响，保证文本内容的清晰和一致。  中期会设计提示语以及评分标准，获得最终的综合评分，结构完整性、逻辑清晰度、语言连贯性、内容独特性和创新性、参考文献规范性、课程知识掌握度六维度小分，以及打分原因和修改意见。  后期将针对不同模型进行测试以及提示语进一步优化，以及根据毕业设计论文此类专业性强的长文本调整标准以及提示语。并且完成前后端设计、优化，合并成完整的应用。  参考文献：   1. LongLLMLingua: Accelerating and Enhancing LLMs in Long Context Scenarios via Prompt Compression (https://doi.org/10.48550/arXiv.2310.06839) 2. Are Large Language Models Good Essay Graders? (https://doi.org/10.48550/arXiv.2409.13120) 3. Automatic Essay Multi-dimensional Scoring with Fine-tuning and Multiple Regression (https://doi.org/10.48550/arXiv.2406.01198) | | | | | | | |
| 发出任务书日期：2024.11　　　　　　　　　完成期限：2025.05  刘江电子签名  指导教师（签名）：  2024年11月30日 | | | | | | | |
| 系/研究中心毕业设计（论文）工作小组审定意见：  　　　　　 主任（签名）：    　 年　　月　　日 | | | | | | | |

备注：

1.论文题目须与论文封面题目一致。

2.任务书一经审定，不得随意更改，如因特殊情况需变更，须经系/研究中心毕业设计（论文）工作小组同意。