### ****基于大语言模型的学术文档摘要系统的设计与实现****

#### ****本科毕业设计开题报告汇报****

#### ****一、研究背景及意义****

随着人工智能技术的迅猛发展，大语言模型（Large Language Model, LLM）在自然语言处理领域的应用日益广泛。学术文档摘要是科研工作中至关重要的环节，但传统方法往往存在人工工作量大、耗时长等问题。大语言模型的引入为自动摘要生成提供了新的可能性。目前，国内外主要研究方向包括：

1. **基于规则的方法**：通过预设规则和知识图谱进行文本解析，效率过低且信息易出现不完整和不流畅等情况。
2. **基于机器学习的方法**：采用深度学习模型如Seq2Seq和Transformer辅助生成摘要，提升摘要生成质量，但仍存在许多问题。
3. **基于大语言模型的方法**：利用GPT-4、T5等强大语言模型生成高质量摘要，在多语言和大规模数据处理中展现速度、质量等优势。

尽管已有研究取得一定进展，但仍存在摘要质量参差不齐、领域适配性不足、页面使用不够便捷等问题。尤其在科研文献爆炸式增长的背景下，现有的文献管理工具在智能化和效率方面尚未完全满足需求。因此，本课题的研究具有重要的现实意义，旨在设计一套高效、便捷的学术文档摘要系统，以提升科研工作者的文献获取效率。

#### ****二、研究内容与技术路线****

本课题主要研究以下内容（预期要完成的内容）：

1. **需求分析**：明确科研人员对摘要系统的核心需求，包括摘要生成、论文分类、关键词搜索等功能。
2. **大语言模型API调用**：选择适合的LLM（如GPT-4），通过API实现摘要自动生成，并进行实验验证可行性。
3. **原型系统开发**：搭建基于前后端架构的学术摘要系统，实现文献上传下载、摘要生成、手动分类和关键词搜索与高亮等核心功能。
4. **系统测试与优化**：对系统进行功能测试、用户体验优化，提高摘要质量与交互便捷性。

**技术路线：**

* **前端开发**：采用Vue.js或React构建用户界面，支持文档上传、摘要查看等功能。
* **后端开发**：使用Java/Spring Boot搭建后端服务，连接LLM API，实现摘要生成。
* **数据库管理**：利用MySQL数据库存储用户数据、文档信息，并根据需求支持云存储。

#### ****三、系统架构设计****

系统整体采用前后端分离架构，包含以下主要模块（预期要完成的内容详细模块）：

1. **文档上传与解析**：支持PDF/Word等格式的学术文档上传与内容提取。
2. **摘要生成模块**：通过LLM API调用，实现自动生成高质量摘要。
3. **分类与关键词搜索**：根据论文内容进行手动分类，并支持关键词搜索与高亮显示。
4. **用户交互界面**：优化用户体验，提供更便捷美观的功能。

#### ****四、实验与评估****

为了验证系统的有效性，将采用以下方法进行评估：

1. **功能测试**：确保各模块功能正常运作，包括摘要生成、分类、搜索等。
2. **摘要质量评估**：采用一些自动评测指标，并结合使用反馈进行质量优化。
3. **用户体验分析**：邀请人员试用，收集使用反馈并改进系统交互设计。

#### ****五、研究计划与进度安排（预期时间安排）****

| **阶段** | **任务** | **时间** |
| --- | --- | --- |
| **1** | 课题调研、需求分析、系统设计 | 2025年1月 |
| **2** | 前后端系统开发与测试 | 2025年2-4月 |
| **3** | 系统优化、功能完善、用户测试 | 2025年5月 |
| **4** | 论文撰写与答辩准备 | 2025年5-6月 |

#### ****六、总结****

本研究旨在利用大语言模型构建一套高效、便捷的学术文档摘要系统，以优化科研人员的文献管理和阅读效率。系统将在确保摘要质量的同时，提升交互体验和功能扩展性。后续工作将围绕系统优化与用户测试展开，确保最终成果的实用性与可推广性。

以上是本次开题报告的主要内容，谢谢大家！