**KẾ HOẠCH THỰC HIỆN**

**SINH VIÊN:**

**1. Trần Thị Thùy Dương- 110121267**

**2. Mã Đại Phú- 110121266**

**3. Nguyễn Hoàng Nhựt-110121269**

**1. Nội dung: Xây dựng hệ thống thư viện số thông minh**

**Mục tiêu chung:** Xây dựng hệ thống thư viện số thông minh cho phéptra cứu tài liệu dạng bản scan.

**Mục tiêu cụ thể:**

* Xây dựng module cho phép trích xuất nội dung file scan.
* Thiết kế cơ sở dữ liệu lưu trữ tập tin văn bản (file PDF) dưới dạng vector.
* Xây dựng chức năng tìm kiếm tài liệu đã lưu trữ có trích dẫn nguồn.

|  |  |
| --- | --- |
| **Thời gian**  **thực hiện** | **Nội dung công việc**  *(mô tả chi tiết công việc cần thực hiện)* |
| **Tuần 1:** Xây dựng module cho phép trích xuất nội dung file scan. | Khảo sát công nghệ/ cách làm hiện có |
| Lựa chọn/ phát triển kỹ thuật trích xuất nội dung file PDF |
| Kiểm thử đánh giá độ chính xác của sản phẩm/giải pháp |
| Điều chỉnh/ cải thiện |
| **Báo cáo kết quả tuần 1** |
| **Tuần 2:** Thiết kế cơ sở dữ liệu lưu trữ tập tin văn bản (file PDF) dưới dạng vector | Nghiên cứu QDRANT, Metadata ( Tên tài liệu, Trang, Đường dẫn lưu trữ) |
| Tạo giao diện thêm tài liệu mới, chỉnh sửa, xóa tài liệu hiện có. |
| Nhập 10 file PDF (sách, tạp chí) |
| **Báo cáo kết quả tuần 2** |
| **Tuần 3:** Xây dựng chức năng tìm kiếm tài liệu đã lưu trữ có trích dẫn nguồn | Xây dựng giao diện tra cứu (Streamlit)  Thiết kế giao diện Streamlit  Xây dựng form nhập liệu  Hiển thị kết quả tìm kiếm có trích dẫn nguồn.  **Báo cáo kết quả tuần 3** |
| **Tuần 4:** Tích hợp LLM Framework | Kết nối Ollama với Streamlit Gọi model Deepseek-R1 hoặc Gemini Free API |
| **Báo cáo kết quả tuần 4** |
| **Tuần 5:** Tích hợp tìm kiếm tài liệu, hoàn thiện cơ chế RAG | Truy xuất vector từ Qdrant  Thực hiện tìm kiếm từ giao diện Streamlit Hiển thị kết quả có liên quan đến từ khóa tìm kiếm (Tên tài liệu, trang, dường dẫn lưu trữ)  Kiểm thử giao diện, kiểm tra mức độ chính xác của kết quả tìm kiếm |
| **Báo cáo kết quả tuần 5** |
| **Tuần 6. Viết báo cáo** | Test tải lớn, đảm bảo hệ thống ổn định  Tổng hợp báo cáo & chuẩn bị tài liệu bàn giao |
| **Báo cáo kết quả tuần 6** |