**Tích hợp và phát triển công cụ tạo câu hỏi " Quiz Generate"**

**< AC >**



**Hà Nội, tháng 12 năm 2024.**

**PHÊ DUYỆT CỦA TRƯỞNG BỘ PHẬN**

|  |  |  |
| --- | --- | --- |
|  | **NGƯỜI SOẠN THẢO** | **NGƯỜI KIỂM TRA** |
| Chữ ký |  |  |
| Họ tên |  |  |
| Chức vụ |  |  |
| Ngày |  |  |

**Bảng theo dõi nội dung sửa đổi tài liệu**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lần sửa đổi | Ngày sửa đổi | Vị trí sửa đổi | Nội dung sửa đổi | Ghi chú |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**Table of content**

[I. Tổng Quan 4](#_Toc187154094)

[1. Mục đích 4](#_Toc187154095)

[2. Mục tiêu 4](#_Toc187154096)

[3. Phạm vi 4](#_Toc187154097)

[3. Abbreviations 5](#_Toc187154098)

[II. Công nghệ và thư viện sử dụng 5](#_Toc187154099)

[1. Thư viện Python 5](#_Toc187154100)

[2. API và Framework 5](#_Toc187154101)

[3. Quy trình thực hiện 5](#_Toc187154102)

[3.1 Tiền xử lý dữ liệu 5](#_Toc187154103)

[3.2 Xử lý văn bản 6](#_Toc187154104)

[3.3 Tạo câu hỏi 6](#_Toc187154105)

[3.4 Xuất định dạng đầu ra 6](#_Toc187154106)

[3.5 Tích hợp API 6](#_Toc187154107)

[3.6 Workflow Chatbot 6](#_Toc187154108)

[4. Kết luận 6](#_Toc187154109)

# 

# 

# **I.** **Tổng Quan**

## **1. Mục đích**

Dự án " Quiz Generate" nhằm phát triển một công cụ tự động hóa việc tạo câu hỏi tự luận dựa trên nội dung bài học. Công cụ này hỗ trợ các định dạng tài liệu phổ biến như PDF, Word, và PowerPoint, giúp giáo viên và nhà đào tạo dễ dàng đánh giá mức độ tiếp thu của học sinh sau mỗi bài giảng.

## **2. Mục tiêu**

* Tự động trích xuất nội dung từ tài liệu học tập và chuyển đổi thành câu hỏi tự luận.
* Cải thiện chất lượng câu hỏi thông qua AI Generative và tối ưu hóa thuật toán tạo câu hỏi.
* Hỗ trợ xuất câu hỏi ra nhiều định dạng khác nhau nhằm đáp ứng đa dạng nhu cầu sử dụng.
* Xây dựng API mạnh mẽ và dễ tích hợp vào các hệ thống giáo dục hiện có.
* Cung cấp chatbot giúp người dùng dễ dàng tương tác và tạo câu hỏi nhanh chóng. Dự án "End Lesson Quiz Generate" nhằm phát triển một công cụ tự động hóa việc tạo câu hỏi tự luận dựa trên nội dung bài học. Công cụ này hỗ trợ các định dạng tài liệu phổ biến như PDF, Word, và PowerPoint, giúp giáo viên và nhà đào tạo dễ dàng đánh giá mức độ tiếp thu của học sinh sau mỗi bài giảng.

## **3. Phạm vi**

* Trích xuất nội dung từ tài liệu dạng .pdf, .docx, .pptx.
* Sử dụng mô hình AI để tóm tắt nội dung và tạo câu hỏi tự luận.
* Xuất câu hỏi ra các định dạng PDF, Markdown, HTML hoặc tạo liên kết Kahoot.
* Xây dựng API hỗ trợ tự động hóa quá trình tạo câu hỏi.

## **3. Abbreviations**

|  |  |  |
| --- | --- | --- |
| STT | Abbreviations | Explain |
|  |  |  |

# **II. Công nghệ và thư viện sử dụng**

## **1. Thư viện Python**

* **LangChain**: Xây dựng pipeline xử lý dữ liệu văn bản và tích hợp LLM.
* **LangChain Google Generative AI**: Mô hình AI sử dụng Gemini để tạo câu hỏi.
* **PyPDFLoader**: Xử lý file PDF.
* **python-docx**: Đọc và xử lý file Word.
* **python-pptx**: Trích xuất nội dung từ PowerPoint.
* **FPDF**: Xuất câu hỏi ra file PDF.
* **FastAPI**: Xây dựng API phục vụ cho hệ thống tạo câu hỏi.

## **2. API và Framework**

* **Google Generative AI (Gemini)**: Hỗ trợ sinh câu hỏi tự luận thông qua AI.
* **FastAPI**: Cung cấp endpoint cho người dùng tải tài liệu lên và nhận file câu hỏi đầu ra.

## **3. Quy trình thực hiện**

### 3.1 Tiền xử lý dữ liệu

* Kiểm tra và phân loại file tải lên.
* Trích xuất nội dung từ các định dạng file hỗ trợ.
* Xử lý lỗi và kiểm tra dữ liệu trống.

### 3.2 Xử lý văn bản

* Chia nhỏ nội dung bằng RecursiveCharacterTextSplitter.
* Tóm tắt nội dung chính để làm cơ sở tạo câu hỏi.

### 3.3 Tạo câu hỏi

* Sử dụng PromptTemplate để hướng dẫn mô hình AI tạo câu hỏi.
* Sinh câu hỏi và câu trả lời mẫu.
* Kiểm tra chất lượng câu hỏi.

### 3.4 Xuất định dạng đầu ra

* Chuyển đổi câu hỏi sang định dạng PDF.
* Hỗ trợ xuất Markdown, HTML hoặc liên kết Kahoot.

### 3.5 Tích hợp API

* Xây dựng endpoint /generate\_text cho phép tải file lên.
* Xử lý yêu cầu và trả về file PDF chứa câu hỏi.
* Xử lý lỗi và phản hồi chính xác đến người dùng.

### Workflow Chatbot

* Tích hợp chatbot sử dụng LangChain để giao tiếp và tạo câu hỏi linh hoạt.
* Cho phép người dùng nhập liệu trực tiếp để sinh câu hỏi theo yêu cầu

## **4. Kết luận**

Công cụ " Quiz Generate" giúp tự động hóa quá trình tạo câu hỏi, tiết kiệm thời gian và nâng cao hiệu quả đánh giá học tập. Hệ thống có thể được mở rộng và cải tiến để phục vụ nhiều lĩnh vực giáo dục khác nhau.