

Đại học quốc gia thành phố Hồ Chí Minh  
Trường đại học Bách khoa  
Khoa Khoa học và kỹ thuật máy tính



Đồ án chuyên ngành  
**BÁO CÁO**

**XÂY DỰNG HỆ THỐNG GIÁO DỤC  
THÔNG MINH DỰA TRÊN MÔ HÌNH  
NGÔN NGỮ LỚN**

Ngành: Khoa học máy tính

**THESIS COMMITTEE:**

- GVHD 1:** Assoc. Prof. Võ Thị Ngọc Châu, PhD  
**GVHD 2:** Assoc. Prof. Nguyễn Hứa Phùng, PhD

—oo—

**Sinh viên 1:** Đỗ Phương Nam (2114111)

**Sinh viên 2:** Phan Phạm Thi (2114857)

**Sinh viên 3:** Nguyễn Trường Tuấn Anh (2112796)

Ho Chi Minh City, June 2024



## **Lời cam kết**

Bài báo cáo này được thực hiện với mục tiêu nghiên cứu và phát triển một hệ thống học tập trực tuyến thông minh, đáp ứng nhu cầu học tập cá nhân hóa và tối ưu hóa quá trình học của người dùng, đặc biệt là sinh viên ngành Khoa học máy tính thuộc khoa Khoa học & Kỹ thuật máy tính, trường Đại học Bách Khoa - Đại học Quốc gia Thành phố Hồ Chí Minh. Trong quá trình thực hiện, chúng tôi đã gặp không ít thử thách và khó khăn, từ việc nghiên cứu các mô hình ngôn ngữ lớn (LLM) cho đến việc triển khai các tính năng AI hỗ trợ trong hệ thống. Tuy nhiên, với sự nỗ lực và hỗ trợ từ nhiều phía, chúng tôi đã và đang hoàn thành được mục tiêu nghiên cứu và phát triển hệ thống như mong muốn.

Mặc dù vậy, chúng tôi cũng nhận thức được rằng trong quá trình thực hiện vẫn còn những hạn chế và khó khăn, đặc biệt trong việc tối ưu hóa một số tính năng AI và việc tạo ra các bài tập tự động. Tuy nhiên, những thử thách này sẽ là nền tảng để chúng tôi tiếp tục cải tiến hệ thống trong các giai đoạn phát triển tiếp theo. Chúng tôi hy vọng rằng những đóng góp của chúng tôi sẽ mang lại giá trị thực tế cho việc học tập và giảng dạy lập trình, đồng thời mở ra cơ hội ứng dụng rộng rãi hơn trong môi trường giáo dục.

Chúng tôi xin cam kết rằng báo cáo này là công trình nghiên cứu độc lập của nhóm tôi. Tất cả các kết quả nghiên cứu, dữ liệu thu thập và phân tích đều được thực hiện một cách trung thực và không sao chép từ bất kỳ công trình nghiên cứu nào trước đó. Mọi nguồn tài liệu tham khảo đều được trích dẫn đầy đủ và rõ ràng trong báo cáo này. Nhóm tôi cam kết rằng những kết quả nghiên cứu trong báo cáo này phản ánh đúng những nỗ lực và công sức mà chúng tôi đã bỏ ra trong suốt quá trình thực hiện. Bất kỳ sự trợ giúp hay đóng góp nào từ bên ngoài đều được ghi nhận đầy đủ và không có sự gian lận hay sao chép.

**Tác giả**

*Đỗ Phương Nam*

*Phan Phạm Thi*

*Nguyễn Trường Tuấn Anh*



## **Lời cảm ơn**

Để hoàn thành bài báo cáo này, chúng tôi xin gửi lời cảm ơn sâu sắc đến các thầy cô, bạn bè và những người đã đồng hành cùng tôi trong suốt quá trình nghiên cứu và phát triển hệ thống. Đặc biệt, chúng tôi xin gửi lời cảm ơn đến giáo viên hướng dẫn, người đã cung cấp cho chúng tôi những kiến thức quý báu, giúp chúng tôi hoàn thiện công trình nghiên cứu này. Sự chỉ dẫn tận tình của thầy cô đã giúp chúng tôi hiểu rõ hơn về các phương pháp và công nghệ tiên tiến trong lĩnh vực học máy và trí tuệ nhân tạo cũng như ứng dụng chúng vào trong lĩnh vực phát triển phần mềm - công nghệ phần mềm.

**Tác giả**

*Đỗ Phượng Nam*

*Phan Phạm Thị*

*Nguyễn Trường Tuấn Anh*



## Tóm tắt

Báo cáo này trình bày thiết kế và phát triển một hệ thống học tập trực tuyến thông minh dành cho sinh viên ngành Khoa học máy tính và những người quan tâm đến việc học lập trình. Mục tiêu chính của hệ thống là cung cấp trải nghiệm học tập cá nhân hóa thông qua việc ứng dụng các mô hình ngôn ngữ lớn (LLM), tối ưu hóa quá trình học tập để trở nên linh hoạt, hiệu quả và phù hợp với nhu cầu cụ thể của từng người học.

Hệ thống tập trung vào việc tăng cường sự tương tác giữa người học và trí tuệ nhân tạo (AI), tạo ra một môi trường học tập thông minh. Các tính năng chính bao gồm quản lý khóa học, lộ trình học tập cá nhân hóa, gia sư AI hỗ trợ lập trình, tạo bài tập tự động và theo dõi tiến độ học tập. Hệ thống hỗ trợ giảng viên trong việc tạo và quản lý khóa học, đồng thời tự động tạo bài tập và đánh giá, với khả năng điều chỉnh độ khó phù hợp với tiến độ học của người học.

Một tính năng nổi bật của hệ thống là gia sư AI, hỗ trợ người học trong việc hiểu mã nguồn, đưa ra gợi ý tối ưu hóa mã, hỗ trợ sửa lỗi và cung cấp phản hồi thời gian thực về chất lượng mã lệnh. Ngoài ra, hệ thống còn tự động tạo bài tập, bao gồm các câu hỏi trắc nghiệm và bài tập lập trình, được điều chỉnh theo tiến độ học và nhu cầu của người dùng.

Hệ thống học tập trực tuyến thông minh này nhằm tối ưu hóa trải nghiệm học tập, làm cho quá trình học trở nên cá nhân hóa, thích ứng và hiệu quả hơn, từ đó giúp người học nâng cao kỹ năng lập trình và tiếp cận kiến thức một cách linh hoạt, hiệu quả.

# Mục lục

<b>1</b>	<b>Giới thiệu về đề tài</b>	<b>13</b>
1.1	Động lực . . . . .	13
1.2	Mục tiêu . . . . .	14
1.3	Phạm vi . . . . .	14
1.4	Ý nghĩa của đề tài . . . . .	15
1.4.1	Ý nghĩa của dự án dưới góc nhìn thực tiễn . . . . .	15
1.4.2	Ý nghĩa của dự án theo góc nhìn khoa học . . . . .	15
1.5	Cấu trúc báo cáo . . . . .	15
<b>2</b>	<b>Kiến thức nền tảng</b>	<b>17</b>
2.1	Kiến thức nền tảng về giáo dục và giáo dục thông minh . . . . .	17
2.1.1	Khái quát về giáo dục . . . . .	17
2.1.2	Khái quát về giáo dục thông minh . . . . .	18
2.2	Mô hình ngôn ngữ . . . . .	18
2.3	Các thuật ngữ kỹ thuật liên quan đến giáo dục thông minh . . . . .	19
2.3.1	Đồ thị tri thức (Knowledge graph) . . . . .	19
2.3.2	Knowledge tracing . . . . .	20
2.4	Kiến trúc công nghệ . . . . .	21
2.4.1	Công cụ xử lý ngôn ngữ tự nhiên - Framework tích hợp LLMs . . . . .	21
2.4.1.1	LangChain . . . . .	21
2.4.1.2	LlamaIndex . . . . .	22
2.4.1.3	Tổng quan . . . . .	23
2.4.2	Hệ quản trị cơ sở dữ liệu đồ thị . . . . .	24
2.4.2.1	Neo4J . . . . .	24



2.4.2.2	ArangoDB . . . . .	24
2.4.2.3	OrientDB . . . . .	24
2.4.2.4	Tổng quan . . . . .	25
2.4.3	Framework xây dựng giao diện người dùng . . . . .	25
2.4.3.1	React . . . . .	25
2.4.3.2	Vue . . . . .	26
2.4.3.3	Angular . . . . .	26
2.4.3.4	Tổng quan . . . . .	27
2.4.4	Framework phát triển API . . . . .	27
2.4.4.1	FastAPI . . . . .	28
2.4.4.2	Django . . . . .	28
2.4.4.3	Tổng quan . . . . .	28
2.5	Tổng kết . . . . .	29
<b>3</b>	<b>Công trình liên quan</b>	<b>30</b>
3.1	Hệ thống đề xuất lộ trình học . . . . .	30
3.2	Ứng dụng LLM vào việc giảng dạy khoa học máy tính, lập trình,... . . . . .	31
3.3	Prompt engineering . . . . .	33
3.4	Retrieval-Augmentation Generation . . . . .	33
3.5	Knowledge tracing/Student Profiling . . . . .	34
<b>4</b>	<b>Hệ thống đề xuất</b>	<b>36</b>
4.1	Mô tả hệ thống . . . . .	36
4.2	Khảo sát yêu cầu . . . . .	37
4.2.1	Người dùng . . . . .	37
4.2.2	Yêu cầu chức năng . . . . .	37
4.2.2.1	User Management . . . . .	37
4.2.2.2	Learning Resource Recommendation System . . . . .	38
4.2.2.3	Exercise Generator . . . . .	38
4.2.2.4	AI Tutor . . . . .	38
4.2.2.5	Progress Tracking . . . . .	39
4.2.2.6	Course Management . . . . .	39



4.2.2.7	Admin Management . . . . .	40
4.2.3	Yêu cầu phi chức năng . . . . .	40
4.2.4	Yêu cầu dữ liệu . . . . .	42
4.2.4.1	Các thực thể và thuộc tính dữ liệu . . . . .	42
4.2.4.2	Mối quan hệ dữ liệu . . . . .	43
4.3	Kết luận . . . . .	44
<b>5</b>	<b>Phân tích và thiết kế</b>	<b>45</b>
5.1	Phân tích . . . . .	45
5.1.1	Business Process Modeling . . . . .	45
5.1.2	Use case diagrams . . . . .	46
5.1.2.1	Course Management . . . . .	47
5.1.2.2	User Management . . . . .	48
5.1.2.3	Learning Resource Recommendation System . . . . .	49
5.1.2.4	Progress Tracking . . . . .	50
5.1.2.5	Exercise Generator . . . . .	52
5.1.2.6	AI Tutor . . . . .	53
5.2	Thiết kế . . . . .	55
5.2.1	Kiến trúc hệ thống . . . . .	55
5.2.1.1	Phân tích kiến trúc hệ thống phổ biến hiện nay . . . . .	55
5.2.1.2	Lựa chọn kiến trúc hệ thống . . . . .	56
5.3	Database . . . . .	60
5.3.1	EER diagram . . . . .	60
5.3.2	Lược đồ CSDL quan hệ ánh xạ . . . . .	62
5.3.3	Class diagram . . . . .	63
5.3.4	Sitemap . . . . .	64
5.3.4.1	Sinh viên . . . . .	64
5.3.4.2	Giảng viên . . . . .	66
5.3.4.3	Manager . . . . .	67
5.4	Thiết kế prompt để xây dựng cơ sở dữ liệu đồ thị và hệ thống đề xuất lộ trình học cho từng khóa học . . . . .	68



5.5	Thiết kế prompt để xây dựng đồ thị tri thức từ nguồn học liệu (bài giảng, module,...) . . . . .	68
5.5.1	Cấu Trúc Đồ Thị . . . . .	69
5.5.2	Thiết kế prompt cho việc đề xuất lộ trình học . . . . .	70
5.6	Thiết kế Wireframe cho giao diện . . . . .	74
5.6.1	Landing Page . . . . .	74
5.6.2	Authentication . . . . .	75
5.6.2.1	Log In Screen . . . . .	75
5.6.2.2	Forgot Password Screen . . . . .	76
5.6.2.3	Reset Password Screen . . . . .	76
5.6.3	Course Management (Student) . . . . .	77
5.6.3.1	Courses List Screen . . . . .	77
5.6.3.2	Detailed Course Screen . . . . .	78
5.6.3.3	Define Goal Modal . . . . .	80
5.6.3.4	Modules . . . . .	80
5.6.3.5	Choices of Learning Styles . . . . .	81
5.6.4	User Profile . . . . .	84
5.6.4.1	Profile Information . . . . .	84
5.6.4.2	Change Password, Avatar . . . . .	84
5.6.5	Dashboard Screen . . . . .	85
5.6.6	Exercise Practice . . . . .	86
5.6.6.1	Code Exercise . . . . .	86
5.6.7	AI Tutor . . . . .	86
5.6.8	Progress Tracking . . . . .	89
<b>6</b>	<b>Hiện thực và Kiểm thử</b> . . . . .	<b>91</b>
6.1	Hiện thực hệ thống . . . . .	91
6.2	Kiến trúc Backend . . . . .	95
6.3	Quá trình triển khai hệ thống . . . . .	96
6.4	Kiểm thử . . . . .	99
6.4.1	Tổng quan về kiểm thử . . . . .	99
6.4.1.1	Loại kiểm thử thực hiện . . . . .	99



6.4.2	Kiểm thử API . . . . .	99
6.4.3	Mô tả chung về các kịch bản kiểm thử . . . . .	99
6.4.3.1	Cấu trúc của một kịch bản kiểm thử . . . . .	100
6.4.3.2	Ví dụ về một kịch bản kiểm thử trong bảng Excel . . . . .	100
6.4.3.3	Mục tiêu của các kịch bản kiểm thử . . . . .	100
6.4.3.4	Kết quả kiểm thử API . . . . .	100
6.4.3.4.a	Kiểm thử phương thức . . . . .	100
6.4.3.4.b	Kết quả tổng hợp của các testcase . . . . .	106
6.4.3.4.c	Hạn chế . . . . .	107
6.5	Thử nghiệm hệ thống với các mô hình ngôn ngữ hiện có . . . . .	108
<b>7</b>	<b>Đánh giá hệ thống</b>	<b>114</b>
7.1	Đánh giá tính năng và hiệu quả của hệ thống . . . . .	114
7.2	Đánh giá các hệ thống liên quan . . . . .	115
7.3	Các vấn đề và khó khăn trong quá trình phát triển . . . . .	116
7.4	Kết chương . . . . .	117
<b>8</b>	<b>Tổng kết</b>	<b>118</b>
8.1	Kết quả đạt được . . . . .	118
8.2	Hướng phát triển . . . . .	118
8.3	Kế hoạch thực hiện Đề án Tốt nghiệp . . . . .	119

# Danh mục hình ảnh

5.1	Activity diagram . . . . .	45
5.2	Use case diagram . . . . .	46
5.3	Detailed use case with id 1: Course Management . . . . .	47
5.4	Detailed use case with id 2: User Management . . . . .	48
5.5	Detailed use case with id 3: Adaptive Learning Path Management of Learner Actor . . . . .	49
5.6	Detailed use case with id 3: Learning Resource Recommendation System of System Actor . . . . .	49
5.7	Detailed use case with id 4: Progress Tracking of Learner Actor . . . . .	50
5.8	Detailed use case with id 4: Progress Tracking of System Actor . . . . .	50
5.9	Detailed use case with id 5: Exercise Generator . . . . .	52
5.10	Detailed use case with id 6: AI Tutor of Learner Actor . . . . .	53
5.11	Detailed use case with id 6: AI Tutor of System Actor . . . . .	53
5.12	Component-based Diagram cho toàn bộ hệ thống . . . . .	59
5.13	Component-based Diagram cho những công nghệ được chọn . . . . .	59
5.14	Sơ đồ EER của hệ thống . . . . .	60
5.15	Lược đồ CSDL quan hệ ánh xạ của hệ thống . . . . .	62
5.16	Class diagram hiện thực của hệ thống . . . . .	63
5.17	Sitemap cho đối tượng Sinh viên . . . . .	64
5.18	Sitemap cho đối tượng Giảng viên . . . . .	66
5.19	Sitemap cho đối tượng Quản trị viên . . . . .	67
5.20	Giải thích chi tiết về các tài nguyên học tập . . . . .	73
5.21	Landing Page . . . . .	74
5.22	Log In Screen . . . . .	75



5.23	Forgot Password Screen . . . . .	76
5.24	Reset Password Screen . . . . .	76
5.25	Courses List Screen . . . . .	77
5.26	Modal Learning Outcomes . . . . .	77
5.27	Detailed Course Screen - Description . . . . .	78
5.28	Detailed Course Page - Lessons . . . . .	78
5.29	Detailed Course Page - Exercises . . . . .	79
5.30	Detailed Course Page - Recommend Lessons . . . . .	79
5.31	Modal để lấy mục tiêu của sinh viên . . . . .	80
5.32	Chi tiết Lesson . . . . .	81
5.33	Modal cho phép sinh viên lựa chọn kiểu học phù hợp . . . . .	81
5.34	Giao diện khi sinh viên chọn học theo cách học làm Quiz với tiến độ tự do . . .	82
5.35	Giao diện khi sinh viên chọn học theo cách học đọc tài liệu với tiến độ tự do . .	83
5.36	Giao diện khi sinh viên chọn học theo cách học làm bài tập Coding với tiến độ tự do . . . . .	83
5.39	Dashboard - Student . . . . .	85
5.40	Giao diện thực hành code . . . . .	86
5.41	Lịch sử submission . . . . .	86
5.42	Các tính năng của AI Tutor . . . . .	87
5.43	Giải thích từng dòng code . . . . .	87
5.44	Hỗ trợ sửa lỗi . . . . .	88
5.45	Cung cấp gợi ý cho bài tập . . . . .	88
5.46	Progress Tracking Page . . . . .	89
6.1	Cấu trúc Frontend . . . . .	92
6.2	Hệ thống Typography sử dụng trong ứng dụng . . . . .	93
6.3	Bảng màu được sử dụng trong ứng dụng . . . . .	94
6.4	Cấu trúc Backend . . . . .	95
6.5	Cấu trúc Cơ sở dữ liệu . . . . .	97
6.6	Ví dụ về một kịch bản kiểm thử . . . . .	100
6.7	Kiểm thử truy xuất các khóa học gần đây . . . . .	101
6.8	Kiểm thử truy xuất các hoạt động gần đây . . . . .	101



6.9 Kiểm thử truy xuất danh sách khóa học của sinh viên . . . . .	102
6.10 Kiểm thử truy xuất chi tiết một khóa học . . . . .	102
6.11 Kiểm thử truy xuất danh sách khóa học được đề xuất . . . . .	103
6.12 Kiểm thử thêm hoạt động gần đây của sinh viên . . . . .	103
6.13 Kiểm thử truy xuất chi tiết một lesson được đề xuất . . . . .	104
6.14 Kiểm thử truy xuất danh sách bài tập quizzes từ module . . . . .	104
6.15 Kiểm thử truy xuất chi tiết một bài tập quiz . . . . .	105
6.16 Kiểm thử bắt đầu làm lại một bài tập quiz . . . . .	105
6.17 Kiểm thử submit một bài tập quiz . . . . .	106
6.18 Kiểm thử truy xuất document từ module . . . . .	106
6.19 Kết quả tổng hợp của các testcase . . . . .	106

# Danh mục bảng biểu

2.1	So sánh giữa LangChain và LlamaIndex . . . . .	23
2.2	So sánh giữa Neo4J, ArangoDB và OrientDB . . . . .	25
2.3	So sánh giữa React, Vue và Angular . . . . .	27
2.4	So sánh giữa FastAPI và Django . . . . .	28
7.1	Đánh giá các hệ thống liên quan . . . . .	115

# Chương 1

## Giới thiệu về đề tài

### 1.1 Động lực

Trong bối cảnh toàn cầu hóa và sự phát triển không ngừng của công nghệ, việc nâng cao chất lượng giáo dục, đặc biệt trong lĩnh vực lập trình, ngày càng trở nên quan trọng. Một trong những thách thức lớn hiện nay là làm thế nào để cá nhân hóa quá trình học lập trình, đáp ứng sự đa dạng trong nhu cầu và khả năng của từng học viên. Các hệ thống giáo dục truyền thống thường gặp khó khăn trong việc tạo ra lộ trình học phù hợp với từng học viên, đặc biệt là trong lập trình, nơi mà khả năng tư duy và phong cách học của mỗi học viên đều khác nhau.

Sự phát triển của công nghệ trí tuệ nhân tạo (AI) và mô hình ngôn ngữ lớn (LLM) đang mở ra cơ hội lớn để xây dựng các hệ thống dạy học lập trình hiệu quả hơn. LLM, với khả năng hiểu và sinh ngôn ngữ tự nhiên, không chỉ giúp học viên giải quyết các vấn đề lập trình nhanh chóng mà còn hỗ trợ cá nhân hóa học tập. LLM có thể cung cấp tài liệu học tập và bài giảng được tùy chỉnh, cũng như hỗ trợ giải quyết bài tập lập trình, từ đó cải thiện hiệu quả học tập.

Tuy nhiên, sự phát triển mạnh mẽ của AI cũng gây ra mối lo ngại về việc sinh viên sử dụng AI để làm bài tập mà không thực sự học hỏi. Điều này đặt ra câu hỏi liệu AI có thể giúp sinh viên tự tìm ra lời giải cho các vấn đề thay vì đưa ra câu trả lời trực tiếp, giúp họ phát triển kỹ năng tư duy giải quyết vấn đề một cách hiệu quả.

Việc chọn giáo dục lập trình làm lĩnh vực nghiên cứu và ứng dụng hệ thống có hai lý do chính:

- Thứ nhất, dạy học lập trình ngày càng trở nên phổ biến trong thời đại internet, giúp học viên phát triển tư duy logic và khả năng giải quyết vấn đề, vì nó yêu cầu phân tích và tư duy hệ thống.
- Thứ hai, lĩnh vực giáo dục lập trình rất gần gũi với ngành khoa học và kỹ thuật máy tính, nơi mà nhóm nghiên cứu đang theo học, tạo ra sự liên kết thực tiễn giữa nghiên cứu và



ứng dụng trong thực tế.

## 1.2 Mục tiêu

Từ động lực đã nêu, mục tiêu của đề tài này gồm các điểm chính như sau:

- **Xây dựng hệ thống học tập cá nhân hóa:** Hệ thống sẽ tập trung vào việc cá nhân hóa trải nghiệm học cho từng học viên, phù hợp với khả năng, trình độ, và nhu cầu học tập riêng biệt. Mỗi học viên có một lộ trình học và cách giải thích riêng, giúp họ tiến bộ một cách hiệu quả nhất. Với khả năng phân tích và hiểu ngữ cảnh của LLM, hệ thống có thể đưa ra các phản hồi phù hợp với nhu cầu học tập của từng học viên.
- **Xác định tính khả thi của việc tích hợp LLM trong giáo dục:** Đánh giá khả năng tích hợp mô hình ngôn ngữ lớn vào hệ thống giáo dục, đặc biệt là trong bối cảnh giáo dục lập trình. Chúng tôi sẽ nghiên cứu tính hiệu quả của LLM trong việc cá nhân hóa và nâng cao khả năng tiếp thu kiến thức của học viên.
- **Giải quyết vấn đề lạm dụng LLM trong giải bài tập:** Hiện nay, sinh viên có thể lợi dụng LLM để giải bài tập lập trình mà không thực sự học. Mục tiêu của đề tài là nghiên cứu xem liệu có thể khiến LLM không trực tiếp đưa ra lời giải cho học sinh mà thay vào đó là cung cấp các câu hỏi gợi mở hoặc hướng dẫn giúp học viên tự giải quyết vấn đề. Điều này nhằm phát triển tư duy giải quyết vấn đề của học viên, thay vì chỉ đưa ra câu trả lời.

## 1.3 Phạm vi

Đề tài này sẽ được thực hiện trong khoảng thời gian 7-8 tháng, bao gồm hai học kỳ tại trường đại học. Phạm vi nghiên cứu chủ yếu tập trung vào lĩnh vực giáo dục công nghệ thông tin, đặc biệt là trong việc giảng dạy lập trình. Cụ thể, hệ thống sẽ được xây dựng để hỗ trợ việc học lập trình cho sinh viên, thông qua việc cá nhân hóa các bài giảng, hướng dẫn và hỗ trợ học tập dựa trên các mô hình ngôn ngữ lớn.

Về mặt công nghệ, nhóm thực hiện đề tài sẽ không xây dựng lại các mô hình ngôn ngữ lớn từ đầu mà sẽ tận dụng các mô hình hiện có, chẳng hạn như GPT, Gemini, hoặc các mô hình tương tự, để khai thác và tùy chỉnh chúng nhằm đáp ứng các yêu cầu cụ thể của hệ thống. Mô hình ngôn ngữ này sẽ được sử dụng để cung cấp các phản hồi tự động, đề xuất bài học, tạo các bài kiểm tra và giúp cá nhân hóa trải nghiệm học lập trình của sinh viên.

Mặc dù hệ thống có thể áp dụng cho các lĩnh vực giáo dục khác, trong phạm vi của đề tài này, mục tiêu chính là tập trung vào việc ứng dụng trong lĩnh vực lập trình và công nghệ thông tin, nhằm hỗ trợ việc giảng dạy và học tập môn học này.



## 1.4 Ý nghĩa của đề tài

### 1.4.1 Ý nghĩa của dự án dưới góc nhìn thực tiễn

- Dự án này mang lại ý nghĩa thực tiễn quan trọng trong việc áp dụng công nghệ AI vào giáo dục, cụ thể là trong việc cá nhân hóa quá trình học tập của học viên. Nhờ vào AI và LLM, hệ thống học tập có thể tạo ra một trải nghiệm học tập hấp dẫn và hiệu quả hơn cho học viên, từ đó giúp tăng động lực học tập.
- AI có thể đóng vai trò như một trợ giảng 1-1, giúp giảm tải áp lực cho giáo viên. AI sẽ hỗ trợ trả lời các câu hỏi cơ bản, đánh giá và chấm điểm tự động cho các bài tập, giúp giáo viên có thể tập trung vào việc hỗ trợ học viên ở mức độ sâu hơn. Điều này không chỉ cải thiện hiệu quả giảng dạy mà còn tạo ra một môi trường học tập linh hoạt và phù hợp với từng cá nhân học viên.

### 1.4.2 Ý nghĩa của dự án theo góc nhìn khoa học

- Dự án này có ý nghĩa khoa học đáng kể trong việc ứng dụng AI và LLM vào giáo dục, đặc biệt là trong giáo dục thông minh. Việc tích hợp LLM sẽ không chỉ mở rộng khả năng của các hệ thống dạy học mà còn đánh giá được hiệu quả của chúng trong việc nâng cao khả năng giải quyết vấn đề và học lập trình.
- Dự án sẽ chứng minh tính ứng dụng của LLM trong việc thực hiện các tác vụ lập trình và giải quyết vấn đề, đồng thời khai thác các lý thuyết giáo dục trong khoa học nhận thức, như lý thuyết tải nhận thức (cognitive load theory) và học thuyết kiến tạo (constructivism). Việc áp dụng LLM vào giáo dục có thể giúp học viên không chỉ tiếp thu kiến thức mà còn phát triển khả năng tự học và tư duy phản biện.

## 1.5 Cấu trúc báo cáo

Báo cáo này được tổ chức thành 8 chương chính, mỗi chương sẽ trình bày một phần quan trọng trong quá trình thực hiện và phát triển hệ thống. Cấu trúc báo cáo như sau:

- **Chương 1 – Giới thiệu:** Cung cấp cái nhìn tổng quan về động cơ, mục tiêu và phạm vi của đề tài, cũng như ý nghĩa của dự án cả trong thực tiễn và trong khoa học. Giới thiệu về cấu trúc báo cáo để hiểu rõ cách thức bố trí và tiếp cận các phần nội dung.
- **Chương 2 – Kiến thức và công nghệ nền tảng:** Trình bày các kiến thức cơ bản và công nghệ nền tảng cần thiết để phát triển hệ thống, bao gồm các khái niệm, lý thuyết và công nghệ được sử dụng trong hệ thống đề xuất.



- **Chương 3 – Công trình liên quan:** Phân tích các nghiên cứu, công trình và hệ thống liên quan đã có trước đây, chỉ ra sự tương đồng và khác biệt với công trình của đề tài này, từ đó xác định các khoảng trống và cơ hội nghiên cứu thêm.
- **Chương 4 – Hệ thống đề xuất:** Mô tả chi tiết hệ thống được đề xuất, bao gồm các mục tiêu, chức năng của hệ thống, người dùng mục tiêu và các yêu cầu kỹ thuật, dữ liệu liên quan.
- **Chương 5 – Phân tích và Thiết kế:** Trình bày quá trình phân tích hệ thống và thiết kế các thành phần của nó, bao gồm mô hình hóa quy trình nghiệp vụ, các giải pháp công nghệ, thiết kế hệ thống, cơ sở dữ liệu và giao diện người dùng.
- **Chương 6 – Hiện thực và kiểm thử:** Mô tả quá trình triển khai thực tế hệ thống và các kiểm thử được thực hiện để đảm bảo hệ thống hoạt động chính xác và hiệu quả.
- **Chương 7 – Đánh giá hệ thống:** Đánh giá hiệu quả và hiệu suất của hệ thống, chỉ ra các ưu điểm, nhược điểm và khả năng cải tiến.
- **Chương 8 – Tổng kết:** Tóm tắt kết quả đạt được, các hạn chế của hệ thống hiện tại và hướng phát triển trong tương lai.

# Chương 2

## Kiến thức nền tảng

### 2.1 Kiến thức nền tảng về giáo dục và giáo dục thông minh

#### 2.1.1 Khái quát về giáo dục

Trong phần này, ta sẽ giới thiệu khái niệm cơ bản về giáo dục. Giáo dục là quá trình truyền đạt và tiếp thu kiến thức, kỹ năng và giá trị, giúp con người phát triển toàn diện về trí tuệ, thể chất và đạo đức. Mục tiêu chính của giáo dục là trang bị cho con người những kiến thức cơ bản cần thiết để hòa nhập vào xã hội, đồng thời phát triển các kỹ năng tư duy, giải quyết vấn đề và sáng tạo.

Bên cạnh đó, giáo dục còn hướng đến việc bồi dưỡng phẩm chất đạo đức, khuyến khích mỗi cá nhân phát triển tiềm năng riêng biệt của mình. Đây chính là nền tảng để mỗi người có thể xây dựng sự nghiệp và cuộc sống thành công.

Trong bối cảnh thế giới không ngừng thay đổi, giáo dục ngày nay không chỉ cung cấp kiến thức mà còn giúp con người thích nghi với những thách thức mới, phát triển khả năng học tập suốt đời và duy trì sự phát triển liên tục qua mọi giai đoạn cuộc sống.

Về lý thuyết giáo dục, có nhiều học thuyết liên quan đến lĩnh vực này, kết hợp từ nhiều lĩnh vực khác nhau như tâm lý học và thần kinh học:

- *Học tập kiến tạo (Constructivism)*: Phương pháp học tập này cho rằng con người tự xây dựng sự hiểu biết và kiến thức của mình thông qua trải nghiệm và việc phản ánh lại các trải nghiệm đó. Thay vì tiếp thu thông tin một cách thụ động, người học chủ động tạo ra ý nghĩa thông qua việc liên kết thông tin mới với kiến thức đã có trước đó.
- *Thuyết tải nhận thức (Cognitive load theory)*: Được phát triển bởi John Sweller vào những năm 1980, lý thuyết này tập trung vào cách bộ não xử lý và lưu trữ thông tin trong quá



trình học tập, đặc biệt là những yếu tố ảnh hưởng đến hiệu quả học tập. Thuyết này nhấn mạnh rằng bộ não con người có giới hạn về lượng thông tin có thể xử lý cùng lúc trong trí nhớ ngắn hạn.

### 2.1.2 Khái quát về giáo dục thông minh

Tuy không có một định nghĩa chính thức và rõ ràng cho thuật ngữ “**giáo dục thông minh**”, trong báo cáo này, thuật ngữ “**thông minh**” được hiểu là cách thức thực hiện mọi hoạt động một cách hiệu quả và hữu ích. Mục tiêu của giáo dục thông minh là bồi dưỡng con người trên nhiều mặt, từ kiến thức đến đạo đức, với phương thức giúp học sinh tiếp thu kiến thức một cách hiệu quả nhất có thể.

Giáo dục thông minh không chỉ liên quan đến việc dạy và học mà còn phải thích ứng với nhu cầu, khả năng, sở thích và điểm mạnh, yếu của từng học sinh. Việc cá nhân hóa quá trình học tập đóng vai trò quan trọng, vì mỗi học sinh có cách tiếp thu khác nhau. Công nghệ trí tuệ nhân tạo (AI) hiện nay đã tạo ra nhiều cơ hội để phát triển giáo dục thông minh, đặc biệt là qua phương pháp học tập cá nhân hóa (personalized learning), nơi mà mỗi học sinh có thể học theo một tốc độ, phong cách và nội dung phù hợp với bản thân.

Học tập cá nhân hóa nhằm tối đa hóa hiệu quả học tập bằng cách điều chỉnh chương trình học và phương pháp giảng dạy để phù hợp với nhu cầu và mục tiêu riêng biệt của mỗi học sinh, thay vì áp dụng một chương trình chung cho tất cả.

## 2.2 Mô hình ngôn ngữ

*Mô hình ngôn ngữ* (*language model*) có thể được hiểu ở nhiều góc độ khác nhau, nhưng nhìn chung, một mô hình ngôn ngữ là một hệ thống được thiết kế để hiểu và tạo ra ngôn ngữ tự nhiên của con người. Trong lĩnh vực machine learning, mô hình ngôn ngữ được định nghĩa rõ ràng hơn là một mô hình xác suất của một ngôn ngữ tự nhiên, có khả năng dự đoán một từ hoặc một chuỗi từ dựa trên những từ đã xuất hiện trước đó.

Trong lịch sử phát triển của lĩnh vực xử lý ngôn ngữ tự nhiên (NLP), đã có nhiều loại mô hình ngôn ngữ khác nhau. Dưới đây là các loại mô hình ngôn ngữ chính:

- **Mô hình ngôn ngữ xác suất:** Là loại mô hình ngôn ngữ là trong đó, mỗi một chuỗi từ ngữ nhất định sẽ được gán một xác suất cụ thể. Bài toán chính mà dạng mô hình này giải quyết đó là tính *khả năng xảy ra* (*likelihood*) của một chuỗi các từ vựng bất kì, câu hỏi đặt ra thường là "Liệu một chuỗi các từ  $w_1, w_2, \dots, w_n$  có khả năng xuất hiện là bao nhiêu?". Một ví dụ điển hình cho dạng mô hình ngôn ngữ xác suất đó là mô hình ngôn ngữ n-gram. Theo mô hình ngôn ngữ này, thì xác suất của một từ xuất hiện phụ thuộc vào các từ trước



đó:

$$P(w_1, w_2, \dots, w_n) = \prod_{i=1}^n P(w_i | w_{i-(n-1)}, \dots, w_{i-1})$$

- **Mô hình ngôn ngữ dựa trên Neural network:** là một loại mô hình ngôn ngữ sử dụng neural network để hiểu và sinh ra ngôn ngữ tự nhiên của con người. Sức mạnh của neural network nằm ở chỗ là nó có thể nhận diện được các pattern phức tạp, do đó, mô hình dạng này giúp thực hiện nhiều nhiệm vụ xử lý ngôn ngữ tự nhiên (NLP) một cách hiệu quả hơn so với các mô hình thống kê truyền thống đã nói ở trên.
- **Mô hình ngôn ngữ lớn:** Dựa trên khái niệm về mô hình ngôn ngữ, ta có khái niệm cho *Mô hình ngôn ngữ lớn (Large language model)*. Mô hình ngôn ngữ lớn là gì? Nó khác gì so với một mô hình ngôn ngữ? Làm sao để biết được một mô hình ngôn ngữ là 'lớn'? Một mô hình ngôn ngữ lớn là một mô hình ngôn ngữ dựa trên *Deep neural network* và được huấn luyện bằng một tập dữ liệu ngôn ngữ khổng lồ. Việc đánh giá một mô hình là to hay nhỏ có nhiều khía cạnh, có thể là quy mô của mô hình hoặc là khả năng. Về quy mô, mô hình ngôn ngữ lớn, đúng như tên gọi, có kích thước rất lớn, với số lượng tham số lên đến con số hàng tỉ, hoặc hàng trăm tỉ. Một ví dụ tiêu biểu có thể kể đến như mô hình GPT-3 của OpenAI có khoảng 175 tỉ tham số [1]. Một mô hình ngôn ngữ lớn có thể thực hiện được nhiều tác vụ hơn, liên quan đến suy luận, giải quyết vấn đề, những thứ mà các mô hình ngôn ngữ truyền thống gần như không thể thực hiện được.

## 2.3 Các thuật ngữ kỹ thuật liên quan đến giáo dục thông minh

### 2.3.1 Đồ thị tri thức (Knowledge graph)

Trước khi nói về thuật ngữ *Đồ thị tri thức*, ta hãy nhìn sơ qua về thuật ngữ *Đồ thị*. Ở đây, nghĩa của từ đồ thị được hiểu theo định nghĩa của nó ở trong *Lý thuyết đồ thị (graph theory)* là một nhánh nghiên cứu của *Toán rời rạc (Discrete mathematics)*: Một đồ thị (graph) là một cấu trúc toán học được sử dụng để mô tả mối quan hệ giữa các đối tượng (được gọi là đỉnh hoặc nút). Các đối tượng này được kết nối với nhau thông qua các cạnh (hay còn gọi là cung) biểu diễn mối quan hệ giữa chúng. Từ định nghĩa của đồ thị, Trong bài báo cáo này, nhóm xin đưa ra định nghĩa của *Đồ thị tri thức (knowledge graph)*: là một dạng thể hiện có cấu trúc, cụ thể hơn là thể hiện bằng đồ thị, trong đó các đỉnh là các thực thể trong thế giới thực như: một đối tượng, một khái niệm, một sự kiện, và các cạnh là mối quan hệ giữa chúng[2].



### **2.3.2 Knowledge tracing**

*Knowledge tracing* là bài toán mô hình hóa kiến thức của học sinh theo thời gian để ta có thể dự đoán chính xác cách học sinh sẽ thực hiện trong các tương tác sau này [3]. Nhờ có *Knowledge tracing*, giảng viên/người giảng dạy/hệ thống có thể kiểm tra được tiến độ học tập cũng như năng lực của mỗi học sinh *một cách hiệu quả hơn* bằng cách theo dõi tương tác của học viên với các học liệu trực tuyến (tài liệu, các câu hỏi, bài kiểm tra,...)[4].



## 2.4 Kiến trúc công nghệ

### 2.4.1 Công cụ xử lý ngôn ngữ tự nhiên - Framework tích hợp LLMs

Trong quá trình phát triển một hệ thống học tập trực tuyến thông minh dựa trên mô hình ngôn ngữ lớn (LLMs), việc lựa chọn công cụ xử lý ngôn ngữ tự nhiên phù hợp là một yếu tố quyết định đến hiệu quả và tính linh hoạt của hệ thống. Với mục tiêu tối ưu hóa trải nghiệm học tập cá nhân hóa, hệ thống cần tích hợp một framework mạnh mẽ để khai thác và xử lý dữ liệu từ nhiều nguồn. Hai công cụ nổi bật trong lĩnh vực này là LangChain và LlamaIndex, đều được thiết kế để hỗ trợ xây dựng các ứng dụng sử dụng LLMs. Báo cáo này sẽ phân tích và so sánh hai framework trên, nhằm xác định công cụ phù hợp nhất cho hệ thống học tập thông minh, dựa trên tính năng, khả năng tích hợp, và yếu tố chi phí.

#### 2.4.1.1 LangChain

Langchain là một thư viện Python được thiết kế để hỗ trợ phát triển các ứng dụng dựa trên mô hình ngôn ngữ lớn (LLMs).

##### Các tính năng chính:

- Tích hợp dễ dàng với nhiều LLMs
  - Dễ dàng tích hợp và chuyển đổi giữa các mô hình ngôn ngữ lớn khác nhau như **GPT-3**, **GPT-4**, **BERT**, hoặc các mô hình tùy chỉnh.
    - \* VD: Sử dụng **GPT-4** cho việc tạo nội dung học tập phức tạp, trong khi sử dụng một mô hình nhẹ hơn như **BERT** cho việc phân tích nhanh câu trả lời của học viên
  - Chuỗi xử lý linh hoạt cho các tác vụ phức tạp
    - \* Cho phép xây dựng các chuỗi xử lý (chains) phức tạp, kết hợp nhiều bước xử lý khác nhau.
      - \* VD: Trong hệ thống gia sư AI, tạo một chuỗi xử lý để phân tích code của học viên, đánh giá, và đưa ra gợi ý cải thiện.
  - Hỗ trợ truy xuất thông tin và tìm kiếm ngữ nghĩa
    - Cung cấp các công cụ để truy xuất thông tin dựa trên ngữ nghĩa, giúp tìm kiếm thông tin liên quan một cách hiệu quả
      - \* VD: Khi học viên đặt câu hỏi, hệ thống có thể tìm kiếm thông tin liên quan từ cơ sở dữ liệu bài giảng bằng cách dùng **FAISS (Facebook AI Similarity Search)** - Một thư viện hỗ trợ sử dụng vector search để tìm kiếm thông tin trong **Knowledge Base**.



- Khả năng xử lý đa dạng các loại dữ liệu
  - Có thể làm việc với nhiều loại dữ liệu khác nhau như văn bản, PDF, hình ảnh (qua tích hợp với các mô hình xử lý hình ảnh).
    - \* VD: Xử lý các tài liệu PDF và chuyển đổi chúng thành nội dung học tập.

#### **2.4.1.2 LlamaIndex**

LlamaIndex là một framework chuyên biệt cho việc xây dựng các ứng dụng AI với khả năng truy xuất dữ liệu mạnh mẽ.

##### **Các tính năng chính:**

- Indexing và truy xuất dữ liệu hiệu quả
  - Chuyên về việc tạo index cho dữ liệu lớn và truy xuất thông tin một cách nhanh chóng và hiệu quả.
    - \* VD: Indexing toàn bộ nội dung khóa học và truy xuất thông tin liên quan.
  - Quá trình Indexing trong LlamaIndex như sau:
    - \* Data Connectors: Đây là các công cụ nhập dữ liệu từ nhiều nguồn khác nhau như API, PDF, cơ sở dữ liệu, hay các ứng dụng bên ngoài (Gmail, Notion, Airtable). Chúng chịu trách nhiệm thu thập dữ liệu và đưa vào hệ thống dưới dạng tài liệu thống nhất.
    - \* Documents/Nodes: Documents là các container chứa dữ liệu từ các nguồn khác nhau, chẳng hạn như từ PDF, API hoặc cơ sở dữ liệu. Nodes là những phần nhỏ của tài liệu, được bổ sung metadata và các mối quan hệ, giúp tăng độ chính xác trong quá trình truy xuất dữ liệu.
    - \* Data Indexes: Sau khi dữ liệu được nhập vào, LlamaIndex sẽ giúp tổ chức dữ liệu thành một định dạng có thể truy xuất. Quá trình này bao gồm việc phân tích, tạo các biểu diễn embedding, và suy luận metadata, tạo thành kho kiến thức để sử dụng sau này.
- Tích hợp với nhiều nguồn dữ liệu
  - LlamaIndex có thể làm việc với nhiều nguồn dữ liệu khác nhau như file local, API, cơ sở dữ liệu.
    - \* VD: Tích hợp dữ liệu từ nhiều nguồn để tạo nội dung khóa học.
- Hỗ trợ cho các truy vấn phức tạp



- LlamaIndex cho phép thực hiện các truy vấn phức tạp, kết hợp nhiều điều kiện và lọc kết quả.
  - \* VD: Tìm kiếm bài học phù hợp dựa trên nhiều tiêu chí.

#### 2.4.1.3 Tổng quan

Tiêu chí	LangChain	LlamaIndex
<b>Overall</b>	Phát triển ứng dụng phức tạp linh hoạt, tích hợp LLM	Tập trung vào tìm kiếm và truy xuất thông tin từ tập dữ liệu lớn một cách nhanh chóng và chính xác
<b>Prompts</b>	Hỗ trợ giao diện tiêu chuẩn cho việc tạo và quản lý prompts, giúp tùy chỉnh và sử dụng lại dễ dàng hơn trên các mô hình khác	Không hỗ trợ chi tiết
<b>Chains</b>	Cung cấp giao diện mạnh mẽ để xây dựng và quản lý chuỗi, cùng với nhiều thành phần có thể tái sử dụng	Không hỗ trợ chuỗi xử lý phức tạp
<b>Agents</b>	Sử dụng LLM để xác định và thực hiện các hành động dựa trên input	Không có cơ chế agent
<b>Data Indexing</b>	Phương pháp indexing qua các chuỗi phức tạp	Lập chỉ mục nhanh chóng các dữ liệu không cấu trúc
<b>Customization</b>	Tùy chỉnh cao cho các workflow và ứng dụng phức tạp	Tùy chỉnh hạn chế, tập trung vào lập chỉ mục và tìm kiếm
<b>Context Retention</b>	Lưu giữ thông tin từ các tương tác trước đó để cho phép các cuộc hội thoại có nhận thức về ngữ cảnh và mạch lạc	Lưu giữ ngữ cảnh cơ bản, phù hợp với nhiệm vụ tìm kiếm
<b>Use Cases</b>	Phù hợp cho các ứng dụng như chatbot, hỗ trợ khách hàng, tạo nội dung phức tạp	Phù hợp cho hệ thống tìm kiếm nội bộ, quản lý kiến thức
<b>Performance</b>	Xử lý tốt các cấu trúc dữ liệu phức tạp	Tối ưu cho tốc độ và độ chính xác trong truy xuất thông tin
<b>Lifecycle Management</b>	Cung cấp bộ đánh giá LangSmith để kiểm tra và gỡ lỗi ứng dụng LLM	Tích hợp công cụ gỡ lỗi và giám sát cho hiệu suất và độ tin cậy

Table 2.1: So sánh giữa LangChain và LlamaIndex

Với những ưu điểm trên, **LangChain** là một lựa chọn tối ưu cho hệ thống vì nó hỗ trợ mạnh mẽ về việc quản lý prompts – một yếu tố quan trọng để tạo ra các hướng dẫn cho các mô hình ngôn ngữ lớn (LLM), cho phép tùy chỉnh và tối ưu chức năng Adaptive Learning Path Management. Ngoài ra, tính năng memory của LangChain giúp lưu giữ và quản lý ngữ cảnh của các cuộc hội thoại trước đó, giúp các tương tác trở nên liền mạch và cá nhân hóa hơn.



## **2.4.2 Hệ quản trị cơ sở dữ liệu đồ thị**

Hệ quản trị cơ sở dữ liệu đồ thị (Graph Database Management Systems - GDBMS) là một loại cơ sở dữ liệu được thiết kế để lưu trữ và truy vấn dữ liệu theo dạng đồ thị, với các nút (nodes), cạnh (edges), và thuộc tính (properties) mô phỏng mối quan hệ giữa các thực thể. Hệ thống này đặc biệt hữu ích trong các trường hợp yêu cầu xử lý các mối quan hệ phức tạp và tìm kiếm các kết nối giữa dữ liệu. Khi lựa chọn hệ quản trị cơ sở dữ liệu đồ thị, sự chọn lựa phụ thuộc vào yêu cầu của ứng dụng, từ khả năng mở rộng, tính linh hoạt đến hiệu suất trong việc xử lý các mối quan hệ phức tạp. Ba công cụ nổi bật trong lĩnh vực này là Neo4J, ArangoDB và OrientDB. Báo cáo này sẽ so sánh ba hệ quản trị cơ sở dữ liệu này dựa trên các tiêu chí về tính năng, hiệu suất và khả năng mở rộng.

### **2.4.2.1 Neo4J**

Neo4J là hệ quản trị cơ sở dữ liệu đồ thị phổ biến nhất, được tối ưu hóa cho việc quản lý và truy vấn các mối quan hệ giữa các đối tượng. Neo4J sử dụng ngôn ngữ truy vấn Cypher và có thể dễ dàng tích hợp vào các ứng dụng phát triển dựa trên đồ thị.

#### **Các tính năng chính:**

- Tối ưu hóa cho các truy vấn đồ thị phức tạp
- Hỗ trợ ACID transactions và các tính năng bảo mật mạnh mẽ
- Tích hợp tốt với các công cụ phân tích dữ liệu và máy học

### **2.4.2.2 ArangoDB**

ArangoDB là một cơ sở dữ liệu đa mô hình, hỗ trợ đồ thị, tài liệu và các kiểu dữ liệu khác. ArangoDB cho phép linh hoạt trong việc lựa chọn mô hình dữ liệu, giúp xây dựng các ứng dụng phức tạp với nhiều loại dữ liệu khác nhau.

#### **Các tính năng chính:**

- Hỗ trợ đa mô hình, bao gồm đồ thị, tài liệu và dữ liệu khóa-giá trị
- Tối ưu cho việc truy vấn đồng thời trên các mô hình dữ liệu khác nhau
- Khả năng mở rộng và hỗ trợ các phân vùng dữ liệu phân tán

### **2.4.2.3 OrientDB**

OrientDB là một cơ sở dữ liệu đồ thị lai, kết hợp các tính năng của đồ thị và cơ sở dữ liệu quan hệ. Nó cung cấp tính linh hoạt và hiệu suất cao trong các ứng dụng cần xử lý mối quan hệ giữa dữ liệu.

**Các tính năng chính:**

- Kết hợp đồ thị và quan hệ trong một hệ thống duy nhất
- Hỗ trợ các mô hình dữ liệu tài liệu, đối tượng và đồ thị
- Các tính năng như truy vấn SQL kết hợp với Cypher (ngôn ngữ của Neo4J)

#### 2.4.2.4 Tổng quan

Tiêu chí	Neo4J	ArangoDB	OrientDB
Mô hình dữ liệu	Đồ thị	Đa mô hình (Đồ thị, Tài liệu, Khóa-Giá trị)	Đồ thị và Quan hệ
Hiệu suất	Tối ưu cho truy vấn đồ thị phức tạp	Tối ưu cho đa mô hình và dữ liệu phân tán	Mạnh mẽ trong các ứng dụng kết hợp đồ thị và quan hệ
Tính mở rộng	Khả năng mở rộng cao với clustering và phân tán dữ liệu	Phân tán dữ liệu linh hoạt với ArangoSearch	Khả năng mở rộng tốt nhưng không mạnh bằng ArangoDB
Tính linh hoạt	Hạn chế ở đồ thị	Hỗ trợ đa mô hình với khả năng linh hoạt cao	Hỗ trợ đồ thị và quan hệ, linh hoạt trong việc lựa chọn mô hình
Tính năng bảo mật	Hỗ trợ các tính năng bảo mật cấp cao	Tích hợp bảo mật và phân quyền linh hoạt	Tính năng bảo mật và phân quyền mạnh mẽ
Khả năng tích hợp	Dễ dàng tích hợp với các công cụ phân tích dữ liệu	Hỗ trợ các công cụ phân tích và tích hợp API mạnh mẽ	Hỗ trợ SQL và tích hợp API tốt

Table 2.2: So sánh giữa Neo4J, ArangoDB và OrientDB

#### 2.4.3 Framework xây dựng giao diện người dùng

Framework xây dựng giao diện người dùng (UI Framework) giúp phát triển các ứng dụng web với giao diện người dùng linh hoạt và dễ dàng tương tác. Việc lựa chọn framework xây dựng giao diện người dùng (UI) là một quyết định quan trọng đối với các ứng dụng web hiện đại. React, Vue, và Angular là ba framework phổ biến nhất hiện nay, mỗi framework có điểm mạnh và hạn chế riêng. Báo cáo này sẽ so sánh ba framework này dựa trên các tiêu chí như tính dễ học, hiệu suất, khả năng mở rộng và sự linh hoạt trong phát triển.

##### 2.4.3.1 React

React là một thư viện JavaScript được phát triển bởi Facebook, chuyên về việc xây dựng UI dựa trên các component tái sử dụng. React nổi bật với khả năng làm việc nhanh và hiệu quả nhờ vào Virtual DOM và hỗ trợ mạnh mẽ từ cộng đồng.

**Các tính năng chính:**



- Virtual DOM giúp tối ưu hiệu suất render và cập nhật giao diện.
- Hệ sinh thái rộng lớn với các thư viện hỗ trợ như React Router, Redux, và React Query.
- Được tối ưu hóa cho các ứng dụng quy mô lớn với khả năng mở rộng mạnh mẽ.
- Yêu cầu kiến thức về JSX (JavaScript XML) và quản lý trạng thái với Redux hoặc Context API.

#### **2.4.3.2 Vue**

Vue là một framework JavaScript nhẹ, dễ học, và được thiết kế để xây dựng các ứng dụng web động. Vue đặc biệt phù hợp với các nhà phát triển mới, với cú pháp đơn giản và dễ sử dụng, nhưng vẫn mạnh mẽ cho các ứng dụng phức tạp.

##### **Các tính năng chính:**

- Cung cấp Vue CLI giúp khởi tạo dự án và xây dựng ứng dụng nhanh chóng.
- Hệ thống component mạnh mẽ và dễ hiểu.
- Vuex cho quản lý trạng thái đơn giản và dễ sử dụng.
- Tích hợp dễ dàng với các dự án hiện tại và khả năng mở rộng cho các ứng dụng lớn.
- Vue Router giúp quản lý routing trong các ứng dụng một cách dễ dàng.

#### **2.4.3.3 Angular**

Angular là một framework full-stack được phát triển bởi Google, được biết đến với khả năng xây dựng các ứng dụng web phức tạp và quy mô lớn. Angular đi kèm với tất cả các công cụ cần thiết để xây dựng một ứng dụng đầy đủ chức năng.

##### **Các tính năng chính:**

- Cung cấp một framework toàn diện, bao gồm công cụ để quản lý routing, form, HTTP requests, và các tính năng bảo mật.
- Tích hợp TypeScript giúp nâng cao tính an toàn của mã nguồn.
- Hệ thống dependency injection và mô hình MVC (Model-View-Controller) giúp tổ chức mã nguồn và quản lý các phần tử trong ứng dụng.
- Phức tạp và yêu cầu học hỏi nhiều hơn so với React và Vue, đặc biệt là khi sử dụng RxJS và các concepts như modules, services và directives.

#### 2.4.3.4 Tổng quan

Tiêu chí	React	Vue	Angular
<b>Dễ học</b>	Dễ học với JSX và các concept như state và lifecycle nhưng yêu cầu làm quen với các thư viện như Redux	Dễ học với cú pháp đơn giản, dễ tiếp cận và tài liệu phong phú	Phức tạp hơn, yêu cầu học các khái niệm như Dependency Injection, RxJS và TypeScript
<b>Hiệu suất</b>	Rất tốt nhờ Virtual DOM và khả năng render nhanh chóng	Cũng sử dụng Virtual DOM nhưng nhẹ và nhanh hơn trong các ứng dụng nhỏ và trung bình	Hiệu suất ổn, nhưng thường chậm hơn trong các ứng dụng lớn do tính phức tạp của framework
<b>Quản lý trạng thái</b>	Quản lý trạng thái mạnh mẽ nhưng phức tạp với Redux hoặc Context API	Quản lý trạng thái đơn giản với Vuex, dễ sử dụng cho cả ứng dụng nhỏ và lớn	Quản lý trạng thái mạnh mẽ với NgRx, nhưng phức tạp hơn Vuex và Redux
<b>Khả năng mở rộng</b>	Phù hợp cho các ứng dụng quy mô lớn và có thể mở rộng dễ dàng nhờ vào hệ sinh thái thư viện mạnh mẽ	Rất linh hoạt và có thể mở rộng cho cả các ứng dụng nhỏ và lớn mà không gặp phải sự phức tạp như Angular	Tốt cho các ứng dụng quy mô lớn với tính cấu trúc chặt chẽ và tổ chức mã nguồn rõ ràng
<b>Cộng đồng</b>	Cộng đồng lớn với nhiều tài nguyên và thư viện hỗ trợ	Cộng đồng đang phát triển, với tài liệu dễ hiểu và hỗ trợ từ cộng đồng ngày càng mạnh mẽ	Cộng đồng lớn, nhưng tài liệu có thể phức tạp và khó tiếp cận hơn đối với người mới bắt đầu
<b>Tính linh hoạt</b>	Linh hoạt cao, nhưng yêu cầu sử dụng các thư viện bổ sung cho các tính năng như routing và quản lý trạng thái	Linh hoạt và dễ dàng tích hợp vào các dự án hiện tại, cung cấp tính năng đầy đủ mà không cần thư viện bên ngoài	Tính linh hoạt thấp hơn so với React và Vue, yêu cầu tuân theo một cấu trúc chặt chẽ và sử dụng các công cụ riêng biệt

Table 2.3: So sánh giữa React, Vue và Angular

Với cú pháp dễ sử dụng, khả năng mở rộng tốt và sự dễ dàng trong việc tích hợp vào dự án hiện tại, nhóm quyết định chọn Vue.js làm framework xây dựng giao diện người dùng cho hệ thống học tập thông minh.

#### 2.4.4 Framework phát triển API

Framework phát triển API là công cụ quan trọng trong việc xây dựng các hệ thống web hiện đại. Việc lựa chọn framework phát triển API là rất quan trọng trong các ứng dụng web và dịch vụ backend. FastAPI và Django đều có những tính năng mạnh mẽ và phù hợp cho việc phát triển API, nhưng mỗi công cụ lại có những ưu điểm riêng. Báo cáo này sẽ so sánh FastAPI và Django dựa trên hiệu suất, tính dễ sử dụng, và khả năng mở rộng.



#### 2.4.4.1 FastAPI

FastAPI là một framework web nhanh chóng được phát triển bằng Python, được tối ưu hóa cho việc xây dựng các API với tính bảo mật cao và dễ dàng mở rộng.

**Các tính năng chính:**

- Tốc độ cao nhờ vào Starlette và Pydantic
- Tự động tạo tài liệu API với OpenAPI
- Hỗ trợ đầy đủ các chuẩn như HTTP, WebSockets, GraphQL, và OAuth2

#### 2.4.4.2 Django

Django là một framework web Python nổi tiếng, cung cấp đầy đủ các tính năng để phát triển ứng dụng web bao gồm API, hệ quản trị cơ sở dữ liệu và bảo mật.

**Các tính năng chính:**

- Cung cấp các công cụ tích hợp như Django REST Framework (DRF) để xây dựng API
- Quản lý cơ sở dữ liệu mạnh mẽ với ORM tích hợp sẵn
- Hệ sinh thái lớn với nhiều thư viện hỗ trợ

#### 2.4.4.3 Tổng quan

Tiêu chí	FastAPI	Django
<b>Hiệu suất</b>	Rất nhanh nhờ vào Starlette và Pydantic	Hiệu suất tốt nhưng không thể so sánh với FastAPI trong các ứng dụng quy mô lớn
<b>Dễ sử dụng</b>	Dễ học, đặc biệt cho người đã quen Python	Được hỗ trợ tốt với tài liệu phong phú và cộng đồng lớn
<b>Tính bảo mật</b>	Tích hợp đầy đủ các công cụ bảo mật hiện đại như OAuth2, JWT	Cung cấp các công cụ bảo mật mạnh mẽ nhưng đòi hỏi cấu hình thêm
<b>Khả năng mở rộng</b>	Rất dễ mở rộng và thích hợp cho các ứng dụng hiện đại	Hỗ trợ mở rộng tốt với kiến trúc có sẵn

**Table 2.4: So sánh giữa FastAPI và Django**

Với hiệu suất cao, tính năng tự động tạo tài liệu API và khả năng dễ dàng tích hợp, nhóm quyết định chọn FastAPI làm framework phát triển API cho hệ thống học tập thông minh.



## 2.5 Tổng kết

Chương 2 đã cung cấp một cái nhìn tổng quan về các khái niệm và lý thuyết nền tảng trong lĩnh vực giáo dục và giáo dục thông minh. Chúng ta đã tìm hiểu về khái niệm giáo dục, vai trò quan trọng của nó trong việc phát triển toàn diện con người và sự thích nghi với môi trường xã hội. Đồng thời, khái niệm giáo dục thông minh đã được làm rõ, nhấn mạnh vào sự kết hợp giữa công nghệ và các phương pháp học tập cá nhân hóa, nhằm nâng cao hiệu quả giảng dạy và học tập.

Các khái niệm về mô hình ngôn ngữ, từ mô hình ngôn ngữ xác suất đến các mô hình ngôn ngữ lớn (LLMs), đã được trình bày chi tiết, làm nền tảng để hiểu rõ hơn về công nghệ ứng dụng trong giáo dục thông minh. Đặc biệt, những thuật ngữ quan trọng như đồ thị tri thức (Knowledge Graph), Knowledge Tracing, và các công nghệ hỗ trợ trong giáo dục như LangChain, LlamaIndex, Neo4J, và VueJS đã được giới thiệu, nhằm làm rõ các công cụ và nền tảng công nghệ cần thiết để phát triển hệ thống học tập thông minh hiện đại.

Kết thúc chương này, ta có thể thấy rằng việc áp dụng công nghệ vào giáo dục không chỉ là một xu hướng mà còn là một yêu cầu thiết yếu để nâng cao chất lượng và hiệu quả học tập. Trong chương tiếp theo, chúng ta sẽ đi sâu vào việc thiết kế và triển khai hệ thống giáo dục thông minh, đồng thời đánh giá các yếu tố ảnh hưởng đến sự thành công của hệ thống này.

# Chương 3

## Công trình liên quan

### 3.1 Hệ thống đề xuất lộ trình học

Trong những năm qua, đã có rất nhiều nghiên cứu và phát triển liên quan đến các hệ thống đề xuất lộ trình học tập. Các hệ thống này không chỉ nhằm cá nhân hóa trải nghiệm học tập mà còn tối ưu hóa quá trình học tập dựa trên dữ liệu từ người học và tài nguyên giáo dục. Từ việc xây dựng đồ thị tri thức để tổ chức các khái niệm học tập, cho đến ứng dụng các mô hình học máy tiên tiến, các công trình này đã đặt nền móng cho việc nâng cao hiệu quả và tính thích nghi của giáo dục thông minh.

- Hệ thống ở bài [5] sử dụng *Đồ thị tri thức* để biểu diễn mối liên hệ giữa các khái niệm (concepts) và các *kết quả học tập* (*learning outcome*). Dựa trên đồ thị tri thức, hệ thống sẽ dùng một giải thuật sinh đường để tạo sinh một lộ trình học tập phù hợp dựa trên đầu vào là *chủ đề bắt đầu* và *chủ đề mục tiêu* của học sinh. Lộ trình học cũng có thể được thay đổi dựa trên năng lực, điểm số,... (Đây là các tham số động - *dynamic parameters*) của học sinh. Đặc biệt, hệ thống liên tục cập nhật độ khó của các tài nguyên học tập (*learning objects – LOs*) cũng như khả năng học tập của học sinh thông qua dữ liệu được thu thập thời gian thực. Các thông số như thời gian học, điểm số đạt được, và số lần thử nghiệm (*number of attempts*) đều được phân tích để điều chỉnh lộ trình sao cho phù hợp nhất với tiến độ học của từng cá nhân. Ngoài ra, hệ thống còn áp dụng các kỹ thuật gợi ý để lựa chọn các tài liệu học tập có mức độ tương thích cao với năng lực hiện tại của học sinh, đảm bảo rằng các tài liệu được gợi ý không chỉ phù hợp về mặt nội dung mà còn tối ưu hóa trải nghiệm học tập.
- Hệ thống FOKE [6] hiện thực một tính năng đề xuất lộ trình học tập cá nhân hóa, tận dụng sức mạnh của các mô hình ngôn ngữ lớn, đồ thị tri thức, và kỹ thuật prompt engineering. Đầu vào của hệ thống bao gồm domain knowledge base, hồ sơ người học, và đặc tả nhiệm



vụ học tập (learning task specification). Các dữ liệu này được xử lý thông qua nhiều mô-đun để tạo ra các gợi ý học tập chính xác, tương tác và có khả năng giải thích. Thành phần cốt lõi của hệ thống là *Rừng tri thức* (*Knowledge Forest - KF*), một phiên bản mở rộng của đồ thị tri thức. Rừng tri thức đại diện cho tri thức miền dưới dạng các cây tri thức (knowledge trees), trong đó mỗi cây thể hiện cấu trúc phân cấp của các khái niệm và mối quan hệ tương ứng. Hồ sơ người học được xây dựng theo cơ chế lập hồ sơ đa chiều (multi-dimensional user profiling), mô hình hóa các đặc điểm tĩnh (thông tin cá nhân, phong cách học tập), hành vi (hoạt động học tập, hiệu suất), và hành trình thời gian (chuỗi kỹ năng đạt được). Dựa trên hai thành phần trên, FOKE kết hợp kỹ thuật *graph embedding* và các gợi ý cấu trúc (structured prompts) để tạo ra các đề xuất lộ trình học tập cá nhân hóa. Đầu ra của hệ thống bao gồm các khuyến nghị về trình tự học tập hoặc tài nguyên học tập, chẳng hạn như khóa học cần hoàn thành, các kỹ năng cần trau dồi, và các tài liệu hỗ trợ.

- Bài báo [7] giới thiệu một phương pháp mới để lập kế hoạch lộ trình học tập cá nhân hóa (*Personalized Learning Path Planning - PLPP*) bằng cách tích hợp các mô hình ngôn ngữ lớn và kỹ thuật thiết kế prompt engineering. Mục tiêu chính của nghiên cứu là cải thiện tính cá nhân hóa, sự thích nghi và tính minh bạch của các hệ thống giáo dục thông minh. Phương pháp đề xuất sử dụng các gợi ý được thiết kế đặc biệt để tích hợp thông tin cụ thể về người học, bao gồm kiến thức nền tảng, mục tiêu học tập và sở thích cá nhân. Các gợi ý này hướng dẫn LLMs, chẳng hạn như GPT-4 và LLama-2-70B, tạo ra các lộ trình học tập cá nhân hóa, phù hợp với nhu cầu và năng lực của từng người học. Ngoài ra, hệ thống hỗ trợ hội thoại đa lượt, cho phép mô hình đặt câu hỏi làm rõ và nhận phản hồi từ người học để chỉnh sửa các đề xuất. Ví dụ, một gợi ý có thể bao gồm: "Dựa trên hiểu biết hiện tại của học sinh về [Chủ đề], đề xuất ba khái niệm tiếp theo mà họ nên học để đạt được trình độ thành thạo trong [Môn học]. Đầu vào của hệ thống bao gồm hồ sơ người học, trạng thái kiến thức hiện tại, và mục tiêu học tập. Đầu ra là các lộ trình học tập được cá nhân hóa, có cấu trúc rõ ràng và được hỗ trợ bởi các giải thích chi tiết về thứ tự học tập được đề xuất. Ví dụ, hệ thống có thể cung cấp lời giải thích như: "Việc học khái niệm [A] trước khái niệm [B] là cần thiết để đảm bảo nền tảng vững chắc trong [Môn học]."

### 3.2 Ứng dụng LLM vào việc giảng dạy khoa học máy tính, lập trình,...

Trong lĩnh vực khoa học máy tính và lập trình, nơi yêu cầu người học không chỉ nắm vững lý thuyết mà còn thành thạo các kỹ năng thực hành, việc tích hợp LLMs đã mang lại nhiều lợi ích nổi bật. Các công cụ dựa trên LLMs không chỉ giúp người học giải quyết các vấn đề kỹ thuật mà còn cung cấp các phản hồi chi tiết, nâng cao khả năng tự học và giải quyết vấn đề. Một số



công trình liên quan đến việc ứng dụng LLM vào việc giảng dạy lập trình có thể kể đến:

- Trong bài báo [8], nhóm nghiên cứu đã tích hợp AI vào khóa học CS50 tại Đại học Harvard với mục tiêu tái hiện tỷ lệ hướng dẫn lý tưởng 1:1 giữa giảng viên và sinh viên. Hệ thống, được gọi là CS50.ai, cung cấp các tính năng hỗ trợ đa dạng như:
  - Giải thích mã nguồn: Sinh viên có thể làm nổi bật một đoạn mã và nhận được giải thích chi tiết bằng ngôn ngữ tự nhiên.
  - Cải thiện phong cách mã nguồn: Công cụ style50 gợi ý các cải tiến phong cách mã lập trình dựa trên hướng dẫn phong cách của khóa học.
  - Trợ lý học tập AI ("CS50 Duck"): Một chatbot tương tác dựa trên GPT-4, hỗ trợ trả lời các câu hỏi học thuật và hành chính liên quan đến khóa học.

Ngoài ra, Hệ thống được thiết kế với "*pedagogical guardrails*" (đường rào sư phạm), đảm bảo rằng AI hướng dẫn thay vì cung cấp trực tiếp đáp án. Kết quả thực nghiệm cho thấy hệ thống tăng cường đáng kể khả năng tự học của sinh viên, đồng thời giảm tải cho giảng viên trong các lớp học lớn và khóa học trực tuyến quy mô lớn (MOOCs).

- Trong bài báo [9], nhóm nghiên cứu triển khai một hệ thống đánh giá lập trình tự động (Automated Programming Assessment System - APAS) Artemis, tích hợp mô hình AI-Tutor dựa trên GPT-3.5 Turbo. Hệ thống được thiết kế để hỗ trợ sinh viên lập trình bằng cách cung cấp phản hồi tức thời và hướng dẫn cải thiện mã nguồn. Tính năng nổi bật của hệ thống bao gồm:
  - Phản hồi theo thời gian thực: Sinh viên có thể nhận phản hồi chi tiết về các lỗi logic, cú pháp và hiệu suất của mã nguồn.
  - Gợi ý cải tiến mã nguồn: AI-Tutor cung cấp các lời khuyên để cải thiện chất lượng mã nguồn, như quản lý bộ nhớ tốt hơn hoặc tối ưu hóa thuật toán.
  - Không tiết lộ đáp án: Hệ thống chỉ cung cấp hướng dẫn, tránh việc tiết lộ giải pháp trực tiếp để đảm bảo người học tự khám phá.

Đầu vào của hệ thống bao gồm bài tập lập trình, mã nguồn hiện tại của sinh viên, và một giải pháp mẫu do giảng viên cung cấp. Đầu ra là các phản hồi chi tiết từ AI về chất lượng mã nguồn của sinh viên và các gợi ý để cải thiện. Nghiên cứu chỉ ra rằng AI-Tutor không chỉ cải thiện chất lượng mã nguồn mà còn giúp sinh viên phát triển tư duy thuật toán. Tuy nhiên, nhóm tác giả cũng lưu ý rằng phản hồi đôi khi còn chung chung và cần cải thiện giao diện để hỗ trợ tương tác tốt hơn.



### 3.3 Prompt engineering

*Prompt engineering* là một hướng đi khá mới do những bước phát triển lớn gần đây của mô hình ngôn ngữ nói chung và mô hình ngôn ngữ lớn nói riêng. Về khái niệm, *Prompt engineering* bao gồm việc thiết kế một cách có chiến lược các hướng dẫn cụ thể cho nhiệm vụ, được gọi là *prompts*, để hướng dẫn đầu ra của mô hình mà không thay đổi các tham số[10]. Một số kĩ thuật tiêu biểu trong prompt engineering có thể kể đến như sau[10]:

- Zero-shot prompting: Mô hình được nhận một tả của công việc mà người dùng muốn nó thực hiện trong prompt nhưng không có dữ liệu được gán nhãn (labeled) để huấn luyện. Sau đó, mô hình tận dụng kiến thức sẵn có của nó để tạo ra các dự đoán dựa trên prompt đã cung cấp để thực hiện tác vụ.
- Few-shot prompting: Khác với Zero-shot prompting, để thực hiện tác vụ, ngoài đặc tả về công việc sẽ được thực hiện, mô hình còn nhận được một số ví dụ cụ thể về công việc mà người dùng muốn mô hình sẽ làm.
- Chain-of-thought prompting[11]: Kĩ thuật này được sử dụng nhằm cải thiện khả năng suy luận của mô hình. Thay vì chỉ đưa ra một câu trả lời trực tiếp, người dùng đưa cho mô hình cung cấp một chuỗi các bước suy nghĩ logic và có hệ thống trước khi đưa ra đáp án cuối cùng.

### 3.4 Retrieval-Augmentation Generation

Một trong những vấn đề của LLM đó là chúng có khả năng trả về những thông tin sai lệch, và Retrieval-Augmentation Generation (hay RAG) là kĩ thuật để hạn chế vấn đề đó, khi LLM sẽ dựa trên một nguồn dữ liệu gọi là *knowledge base* để đưa ra câu trả lời. Có ba giai đoạn chính trong một kĩ thuật RAG[12] thông thường:

- *Indexing*: Raw data từ các nguồn khác nhau (PDF, HTML, Word, Markdown) được làm sạch và trích xuất, sau đó sẽ được chuyển đổi thành định dạng văn bản thuần nhất. Văn bản sẽ được chia ra thành các *chunks* và các chunks này sẽ được biến đổi sang dạng vector và được lưu trong một vector database.
- *Retrieval*: Khi nhận được một câu truy vấn, hệ thống sẽ biến đổi truy vấn thành vector và so sánh độ tương đồng giữa vector biểu diễn cho câu truy vấn và các đoạn văn bản đã được lập chỉ mục. Những đoạn văn bản có độ tương đồng cao nhất sẽ được chọn làm ngữ cảnh mở rộng cho mô hình.



- *Generation:* Truy vấn và các đoạn văn bản được tổng hợp thành một prompt, sau đó mô hình ngôn ngữ lớn sẽ phản hồi cho câu prompt đó. Mô hình có thể dựa trên kiến thức sẵn có hoặc chỉ sử dụng thông tin từ các tài liệu được cung cấp.

Với sự phát triển mạnh mẽ của LLMs trong những năm gần đây, các công trình liên quan đến RAG xuất hiện nhiều. Đầu tiên, đánh chỉ mục (indexing). PDF là một trong những định dạng tài liệu phổ biến nhất trên Internet, và cũng là một trong những nguồn dữ liệu chủ yếu được sử dụng cho việc đánh chỉ mục. Một nghiên cứu được thực hiện ở bài [13] cho thấy, công đoạn trích xuất nội dung từ một tập tin PDF (PDF parsing) có ảnh hưởng rất đáng kể đến chất lượng của việc truy vấn trong RAG. Về nguồn dữ liệu, thay vì chỉ tập trung vào văn bản là chủ yếu, có một số nghiên cứu, kỹ thuật liên quan được phát triển cho các nguồn dữ liệu khác như: Đồ thị tri thức là một dạng dữ liệu có cấu trúc, các dạng dữ liệu giả cấu trúc như PDF[12].

### 3.5 Knowledge tracing/Student Profiling

Với sự phát triển của máy tính và trí tuệ nhân tạo, nhiều công trình liên quan đến knowledge tracing/student profiling cũng đã được đề xuất và đưa ra, với mục tiêu cuối cùng là mô hình được năng lực của người học, từ đó phục vụ cho các bài toán khác trong giáo dục thông minh, đặc biệt là giáo dục thông minh. Một trong những mô hình tiêu biểu và cổ điển nhất trong lĩnh vực này đó là *Bayesian knowledge tracing*, với ý tưởng đó là mô hình hóa kiến thức của học sinh bằng cách sử dụng các biến ngẫu nhiên  $I_j$ , trong đó:

$$I_j = \begin{cases} 1, & \text{Nếu học sinh thành thục kĩ năng thứ } j \\ 0, & \text{ngược lại} \end{cases} \quad (3.1)$$

Ngoài ra, ở trong [14], nhóm tác giả đã xây dựng một mô hình knowledge tracing dựa trên một giả định đó là các khái niệm, mảng kiến thức có liên quan đến nhau về mặt ngữ nghĩa, và nếu như mức độ thành thạo của người học ở một mảng kiến thức nhất định có thể đo được, và người học có một bước tiến ở một kĩ năng, hay một khái niệm nào đó, thì ở các mảng kiến thức liên quan, người học cũng có một bước tiến nhất định tương đối.

Với sự phát triển của Học sâu trong những năm trở lại đây, một nghiên cứu sử dụng một mô hình với tên gọi là *Deep knowledge tracing* [3] để giải quyết bài toán đã nêu, trong đó, ý tưởng của giải pháp đó là sử dụng mô hình *Recurrent Neural Network (RNN)*. Việc sử dụng RNN cho phép ta có thể nhận diện ra được những hình mẫu phức tạp hơn trong việc mô hình hóa kiến thức của học sinh.

Trong những năm trở lại đây, với sự xuất hiện của Mô hình ngôn ngữ lớn, có nhiều nghiên cứu đánh giá và tìm cách ứng dụng Mô hình ngôn ngữ lớn vào trong bài toán Knowledge tracing. Trong [15], (nhóm) tác giả đã đặt ra vấn đề rằng, do đặc tính *black-box* của các mô hình học sâu, nên kết quả của chúng *không thể giải thích được*, và (nhóm) tác giả đã tận dụng hai đặc tính



*suy luận và tạo sinh* của mô hình ngôn ngữ lớn để đưa ra một lời giải thích cho hoạt động đánh giá năng lực của học viên của mô hình.

# Chương 4

## Hệ thống đề xuất

### 4.1 Mô tả hệ thống

Hệ thống học tập trực tuyến thông minh được thiết kế để phục vụ cho sinh viên ngành công nghệ thông tin và những người có nhu cầu học lập trình. Mục tiêu chính của hệ thống là cung cấp trải nghiệm học tập cá nhân hóa thông qua việc ứng dụng các mô hình ngôn ngữ lớn (LLM – Large Language Model), nhằm tối ưu hóa quá trình học tập một cách linh hoạt, hiệu quả và phù hợp với nhu cầu riêng biệt của từng người học.

Hệ thống tập trung vào sự tương tác giữa người học và trí tuệ nhân tạo (AI), tạo ra một môi trường học tập thông minh. Cụ thể, hệ thống hỗ trợ giảng viên dễ dàng tạo, chỉnh sửa và quản lý nội dung học tập. Các chức năng chính của hệ thống bao gồm:

- **Quản lý khóa học:** Giảng viên có thể cung cấp thông tin chi tiết về khóa học, bao gồm tên khóa học, mô tả, mục tiêu học tập, và đối tượng học viên. Các tài liệu học tập có thể được tải lên hoặc liên kết đến các tài nguyên bổ sung.
- **Quản lý nội dung học:** Khóa học được chia thành các module hoặc bài học nhỏ với mục tiêu học tập cụ thể và thời gian hoàn thành dự kiến. Mỗi module hoặc bài học sẽ có mục tiêu học tập rõ ràng để sinh viên biết họ cần đạt được những kỹ năng hoặc kiến thức nào sau khi hoàn thành.
- **Lộ trình học tập cá nhân hóa:** Lộ trình học tập của mỗi sinh viên sẽ được cá nhân hóa thông qua các đề xuất learning item (bao gồm khóa học, module, bài học) dựa trên mục tiêu học tập, nền tảng kiến thức, và hành vi học tập của người dùng (như điểm số, thời gian hoàn thành bài học). Hệ thống sẽ linh hoạt đề xuất các item phù hợp với trình độ hiện tại của người học. Các bài học và module sẽ được điều chỉnh dần theo độ khó tăng lên, giúp người học tiếp cận kiến thức mới một cách hợp lý và tiến bộ liên tục.



- **Gia sư AI hỗ trợ lập trình:** Một trong những tính năng nổi bật của hệ thống là gia sư AI, đóng vai trò như một trợ lý thông minh trong việc giám sát quá trình học của người học. Gia sư AI có khả năng cung cấp các giải thích chi tiết về mã nguồn, gợi ý cách viết mã tối ưu, hỗ trợ sửa lỗi (debug), và đưa ra phản hồi về chất lượng mã lệnh. Tính năng này không chỉ giúp người học nâng cao kỹ năng lập trình mà còn giúp cải thiện hiệu quả học tập thông qua việc tiếp thu kiến thức một cách nhanh chóng và dễ hiểu.
- **Hệ thống tạo bài tập tự động:** Hệ thống có khả năng tự động tạo ra các bài tập lập trình theo chủ đề mà người học mong muốn. Bộ test case cho bài tập sẽ được sinh ra tự động tùy thuộc vào mức độ khó và yêu cầu của bài tập. Người học có thể yêu cầu hệ thống tạo bài tập ở các cấp độ khác nhau (từ cơ bản đến nâng cao) hoặc nhập vào một câu hỏi lập trình cụ thể để hệ thống tạo bộ test case phù hợp.
- **Bài tập sẽ được điều chỉnh linh hoạt:** Bài tập sẽ được điều chỉnh linh hoạt để phù hợp với năng lực và tiến độ học tập của người học. Hệ thống sẽ theo dõi và đánh giá kết quả bài tập dựa trên thời gian giải quyết, độ chính xác, và điểm số, từ đó điều chỉnh lộ trình học tập của người học.
- **Theo dõi tiến độ và báo cáo học tập:** Hệ thống cung cấp công cụ theo dõi tiến độ học tập chi tiết. Sinh viên có thể xem lại thời gian học, số lượng bài tập đã hoàn thành và kết quả của các bài kiểm tra. Hệ thống sẽ cung cấp báo cáo chi tiết về quá trình học tập của sinh viên, bao gồm các bài học đã hoàn thành, kết quả đạt được và các gợi ý từ hệ thống cho bước học tiếp theo. Dựa trên các chỉ số này, hệ thống sẽ đánh giá mức độ tiến bộ của người học và điều chỉnh lộ trình học tập một cách linh hoạt.

## 4.2 Khảo sát yêu cầu

### 4.2.1 Người dùng

Các người dùng liên quan trong hệ thống:

- Học viên.
- Quản trị viên.
- Giảng viên

### 4.2.2 Yêu cầu chức năng

#### 4.2.2.1 User Management

1. **Đăng ký và đăng nhập:** Người dùng có thể tạo tài khoản và đăng nhập bằng thông tin cá nhân hoặc tài khoản mạng xã hội.



2. **Hồ sơ học tập:** Mỗi người dùng có một hồ sơ học tập, lưu trữ thông tin cá nhân, tiến độ học tập, và kết quả các bài kiểm tra/bài tập.
3. **Quản lý lịch sử học tập:** Hệ thống lưu trữ và cho phép người dùng xem lại các bài học, bài tập đã hoàn thành và kết quả đạt được.
4. **Cấu hình cá nhân:** Người dùng có thể cập nhật thông tin cá nhân, độ tuổi.

#### **4.2.2.2 Learning Resource Recommendation System**

1. **Cá nhân hóa đề xuất học tập:** Hệ thống tự động đề xuất các tài nguyên học tập (khóa học, module, bài học, bài tập, project) cho mỗi người dùng dựa trên mục tiêu học tập, độ tuổi, trình độ học vấn, phong cách học tập, và hành vi học tập, lịch sử học tập, mức độ thông thạo của người học trên một khái niệm hoặc một đầu ra học tập cụ thể.
2. **Lựa chọn tài nguyên học tập:** Hệ thống cho phép người dùng tự chọn các learning item dựa trên sở thích hoặc mục tiêu cá nhân, cung cấp sự linh hoạt trong việc tiếp cận nội dung học tập.
3. **Theo dõi phản hồi người dùng:** Hệ thống ghi nhận phản hồi của người dùng về các tài nguyên học tập để cải thiện đề xuất và tối ưu hóa trải nghiệm người học.

#### **4.2.2.3 Exercise Generator**

1. **Tạo bài tập theo mức độ:** Hệ thống tự động sinh ra các bài tập từ cơ bản đến nâng cao phù hợp với khả năng của người dùng.
2. **Tạo bài tập theo yêu cầu:** Người dùng có thể yêu cầu hệ thống tạo bài tập ở các cấp độ khác nhau hoặc nhập vào một câu hỏi lập trình cụ thể.
3. **Tự động điều chỉnh bài tập:** Dựa trên tiến độ học tập, các bài tập sẽ được nâng cấp để phù hợp với mức độ phát triển của người dùng.
4. **Phân tích kết quả bài tập:** Hệ thống đánh giá bài tập dựa trên thời gian giải, độ chính xác, và kết quả điểm số để điều chỉnh lộ trình học.
5. **Tạo bộ test case tự động:** Hệ thống tự động tạo bộ test case phù hợp với yêu cầu của bài tập mà người dùng upload.

#### **4.2.2.4 AI Tutor**

1. Hệ thống có tích hợp một tính năng là Smart Assistant. Khi người dùng tương tác với tài nguyên học tập, sẽ có một hệ thống chatbot hỗ trợ trả lời nội dung liên quan đến tài nguyên học tập mà người dùng đang xem.



2. Ứng với các tài nguyên học tập đòi hỏi người dùng phải thực hành như là các bài tập lập trình, thì hệ thống cũng có tích hợp một số tính năng đi kèm như sau:

- **Hỗ trợ giải thích code:** Người dùng có thể yêu cầu gia sư AI giải thích code hoặc hướng dẫn cách giải quyết các bài tập lập trình.
- **Gợi ý hướng giải:** Khi gặp khó khăn, hệ thống sẽ đưa ra các gợi ý từ gia sư AI nhằm hỗ trợ người dùng hoàn thành bài tập.
- **Đưa ra gợi ý:** Hệ thống có thể hỏi người dùng liệu họ có muốn thay đổi bài tập sang một bài dễ hơn với nội dung tương tự khi gặp vấn đề với bài hiện tại.
- **Tối ưu hóa code:** Gia sư AI cung cấp gợi ý cách viết mã tối ưu hơn.
- **Hỗ trợ sửa lỗi:** Gia sư AI cung cấp hỗ trợ debug khi người dùng gặp lỗi trong code.
- **Đánh giá chất lượng code:** Gia sư AI cung cấp phản hồi kịp thời về chất lượng mã lệnh của người dùng.

#### 4.2.2.5 Progress Tracking

1. **Theo dõi tiến độ học tập:** Hệ thống ghi nhận thời gian học, số lượng bài tập hoàn thành và kết quả các bài kiểm tra.
2. **Đánh giá tiến độ:** Dựa trên các chỉ số như tốc độ giải quyết bài tập và điểm số, hệ thống đánh giá sự tiến bộ của người dùng và điều chỉnh lộ trình học tập.
3. **Báo cáo tiến độ:** Người dùng có thể xem báo cáo chi tiết về quá trình học tập, bao gồm các bài học đã hoàn thành, kết quả đạt được, và gợi ý từ hệ thống cho bước tiếp theo.
4. **Phân tích dữ liệu học tập:** Hệ thống phân tích hành vi học tập của người dùng, bao gồm điểm số, thời gian hoàn thành bài tập, và tần suất yêu cầu hỗ trợ để cải thiện trải nghiệm học tập.

#### 4.2.2.6 Course Management

1. **Tạo khóa học/module:** Giảng viên có thể tạo và thiết kế các khóa học hoặc module học tập mới. Ứng với mỗi một khóa học, giảng viên sẽ phải cung cấp các thông tin về đầu ra học tập (learning outcome), về mô tả của khóa học. Ứng với mỗi khóa học sẽ bao gồm các modules/bài học/projects tương ứng.
2. **Quản lý nội dung:** Giảng viên có thể cập nhật, chỉnh sửa nội dung của khóa học cũng như tài nguyên học tập. Hệ thống cho phép tải lên các tài liệu học tập hoặc liên kết đến các tài nguyên bổ sung.



#### **4.2.2.7 Admin Management**

##### **1. Quản lý người dùng:**

- Quản trị viên có thể xem danh sách người dùng, tìm kiếm, lọc, và sắp xếp theo các tiêu chí như tên, email, trạng thái tài khoản, và tiến độ học tập.
- Quản trị viên có thể khóa/mở khóa tài khoản người dùng hoặc thay đổi quyền truy cập của người dùng (như chuyển từ người dùng bình thường sang giảng viên).

##### **2. Quản lý phản hồi:**

- Xem và xử lý các phản hồi, đánh giá của người dùng về khóa học, giảng viên, và hệ thống.
- Gửi phản hồi cho người dùng về các vấn đề đã được xử lý.

#### **4.2.3 Yêu cầu phi chức năng**

##### **• Hiệu năng**

- Hệ thống phải xử lý yêu cầu truy cập và tương tác của người dùng một cách nhanh chóng và hiệu quả, với thời gian phản hồi không vượt quá 2 giây cho các thao tác thông thường như đăng nhập, xem khóa học, và tải trang.
- Hệ thống phải có khả năng hỗ trợ ít nhất 1000 người dùng đồng thời mà không làm giảm hiệu năng.
- Thời gian phản hồi của các công cụ AI không được quá 7-8s.

##### **• Khả năng mở rộng**

- Hệ thống có khả năng mở rộng khi số lượng người dùng, khóa học, và bài tập tăng lên mà không ảnh hưởng đến chất lượng dịch vụ và thay đổi hệ thống quá nhiều.
- Có thể thêm các chức năng phụ kết hợp với chức năng sẵn có nếu có yêu cầu mở rộng.

##### **• Bảo mật**

- Dữ liệu cá nhân của người dùng và kết quả học tập phải được mã hóa trong quá trình truyền tải và lưu trữ.

##### **• Tính khả dụng**

- Người dùng có thể dễ dàng điều hướng giữa các khóa học, theo dõi tiến trình học tập và thực hiện các bài kiểm tra mà không gặp khó khăn.



- Giao diện người dùng cần trực quan, dễ sử dụng, màu sắc cơ bản.
- Tính tương thích
  - Hệ thống phải tương thích với nhiều trình duyệt phổ biến (Chrome, Firefox, Safari, Edge) và các hệ điều hành khác nhau (Windows, macOS, Linux).



## **4.2.4 Yêu cầu dữ liệu**

### **4.2.4.1 Các thực thể và thuộc tính dữ liệu**

#### **1. User**

- Thực thể: User
- Thuộc tính: ID, username, email, password, role (student, admin, teacher)
  - Student: Goal, Birthyear, Background.
  - Teacher: Organizaiton
  - Admin

#### **2. Course**

- Thực thể: Course
- Thuộc tính: ID, Name, Description, Type, Learning Outcomes, nCredits.

#### **3. LearningPath**

- Thực thể: LearningPath.
- Thuộc tính: ID, Progress, Objective, StartDate, EndDate.

#### **4. Lesson**

- Thực thể: Lesson.
- Thuộc tính: ID, Lesson Outcomes, Description, Name, Documents.

#### **5. Recommend Lesson**

- Thực thể: Recommend Lesson.
- Thuộc tính: Expalin, Content, Progress.

#### **6. Exercise**

- Thực thể: Exercise
- Thuộc tính: ID, Name, Description, Type (Quiz, Code).

#### **7. Module**

- Thực thể: Module
- Thuộc tính: ID, Objective, Title, Access.

#### **8. Quiz**



- Thực thể: Quiz
- Thuộc tính: ID, Status, Score, MaxScore, LLM Response(Name, Difficulty, Question, Answer, Options, Image).

#### 9. Document

- Thực thể: Document
- Thuộc tính: ID, LLM Response.

#### 10. Activity

- Thực thể: Activity
- Thuộc tính: ID, Type, Description, Timestamp.

#### 11. Feedback

- Thực thể: Feedback
- Thuộc tính: ID, Content, Time.

### **4.2.4.2 Mối quan hệ dữ liệu**

- User:
  - Specializes: Teacher, Manager, Student
  - Mối quan hệ với các thực thể:
    - \* Student:
      - Include: 1-N Activity
      - Send: 1-N Feedback
      - Submission: N-N Exercise
      - Join: N-N Course
      - Bookmark: N-N Lesson
    - \* Teacher:
      - Upload: 1-N Course
    - \* Manager:
      - Manage: N-N Course
- Course:
  - Include: 1-N Lesson
  - 1 Student - 1 Course: 1 LearningPath



- Mối quan hệ với các thực thể:
  - \* Include: 1-N Lesson
  - \* 1 Student 1 Course Have 1 LearningPath: Mối quan hệ giữa Student và Course, mỗi học viên chỉ có một LearningPath cho mỗi khóa học.
- Lesson:
  - Specializes: Recommend Lesson
  - Mối quan hệ với các thực thể:
    - \* Include: 1-N Module
- LearningPath:
  - Have: 1-N Recommend Lesson
- Recommend Lesson:
  - Include: 1-N Module
- Module:
  - Generate: 1-N Quiz
  - Generate: 1-1 Document

### **4.3 Kết luận**

Chương 4 đã trình bày chi tiết về hệ thống đề xuất trong môi trường học tập trực tuyến thông minh. Hệ thống cung cấp các công cụ mạnh mẽ giúp người học có trải nghiệm học tập hiệu quả và tối ưu. Các tính năng như tạo bài tập tự động, theo dõi tiến độ học tập, và hỗ trợ từ gia sư AI giúp người học nâng cao khả năng lập trình và tiếp cận kiến thức một cách linh hoạt, hiệu quả.

# Chương 5

## Phân tích và thiết kế

### 5.1 Phân tích

#### 5.1.1 Business Process Modeling

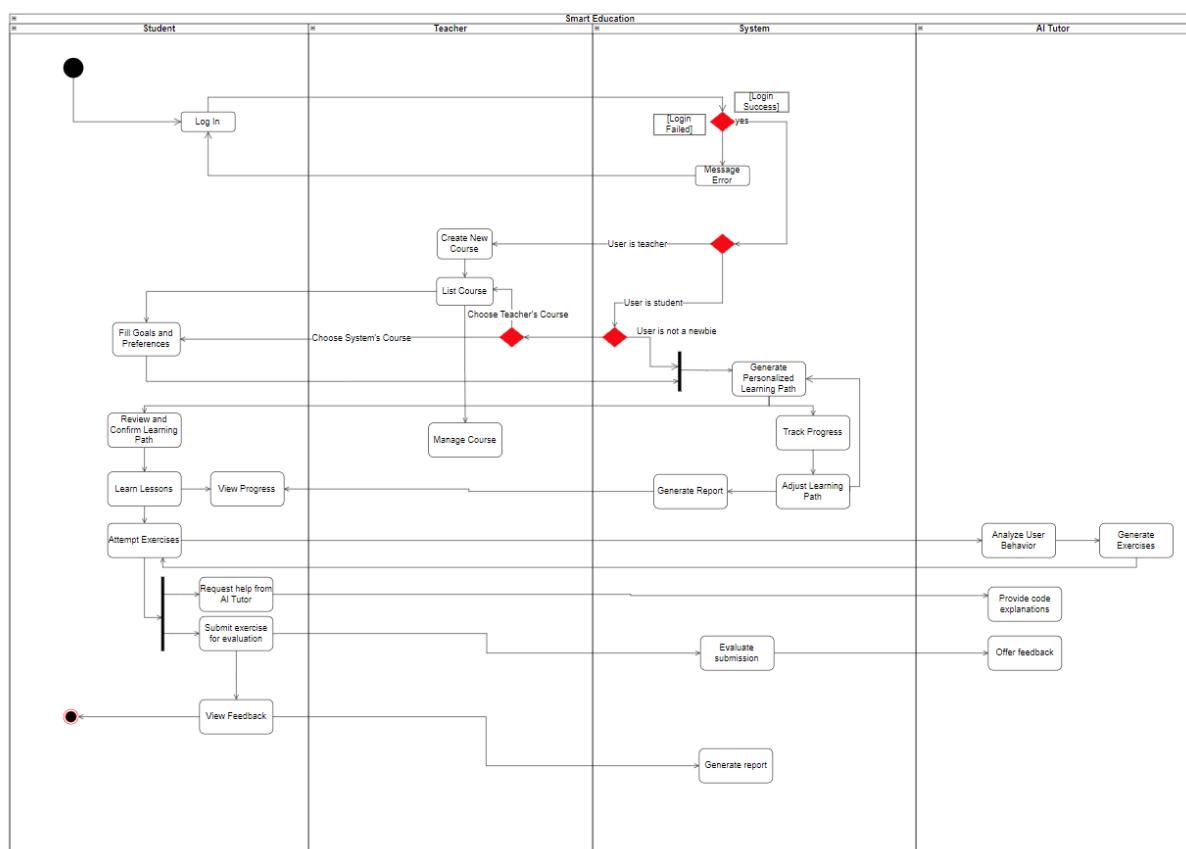


Figure 5.1: Activity diagram

Activity diagram mô tả quy trình tương tác giữa bốn Actors: **Student (Học sinh)**, **Teacher (Giáo viên)**, **System (Hệ thống)**, và **AI Tutor (Gia sư AI)** trong hệ thống giáo dục thông minh.

Học sinh bắt đầu bằng việc đăng nhập và tùy thuộc vào vai trò của họ (giáo viên hoặc học sinh), các hành động khác nhau sẽ diễn ra. Giáo viên có thể tạo khóa học mới hoặc chọn khóa học hiện có để quản lý, trong khi học sinh sẽ xác định mục tiêu học tập và xác nhận lộ trình học do hệ thống đề xuất. Sau đó, học sinh có thể học bài, thực hiện bài tập, và nếu cần có thể yêu cầu trợ giúp từ AI Tutor. Hệ thống cũng tự động theo dõi tiến độ học tập, điều chỉnh lộ trình và đánh giá bài nộp của học sinh. AI Tutor hỗ trợ bằng cách phân tích hành vi, giải thích mã code, và tạo bài tập phù hợp.

### 5.1.2 Use case diagrams

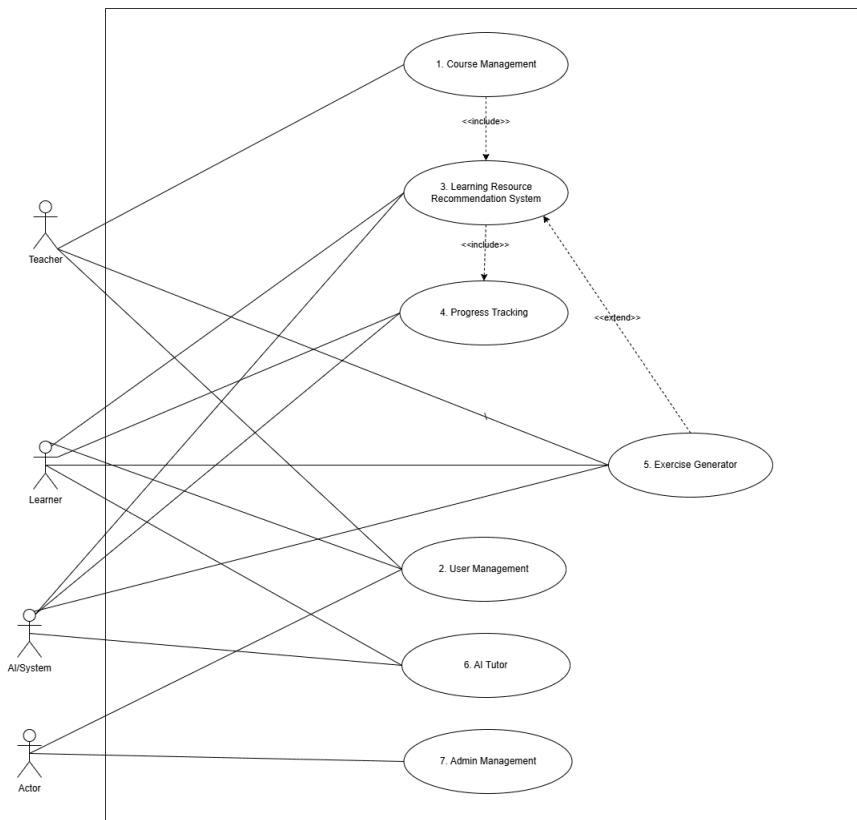


Figure 5.2: Use case diagram

Mô hình use case tổng thể cho hệ thống học tập thông minh bao gồm ba tác nhân chính: Giáo viên, Người học, Quản trị viên và Hệ thống/AI. Các use case bao gồm Quản lý khóa học, Quản lý người dùng, Quản lý lộ trình học tập thích ứng, Theo dõi tiến độ, Tạo bài tập, Quản lý hệ thống cho Quản trị viên và Gia sư AI.

### 5.1.2.1 Course Management

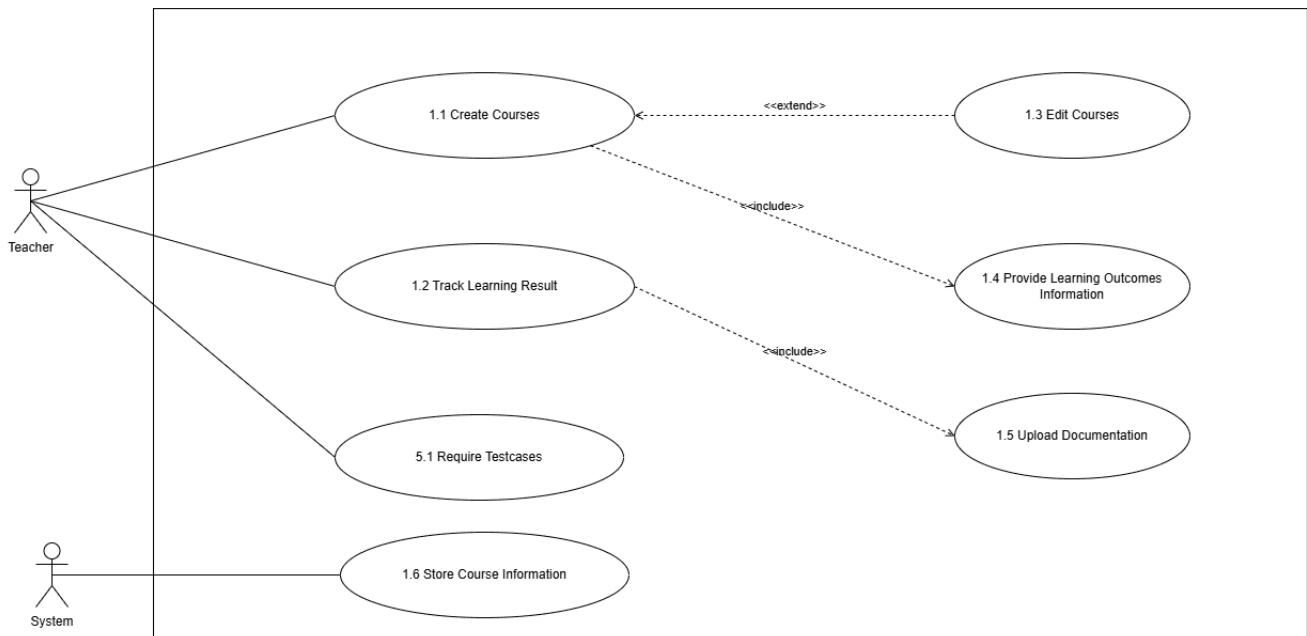
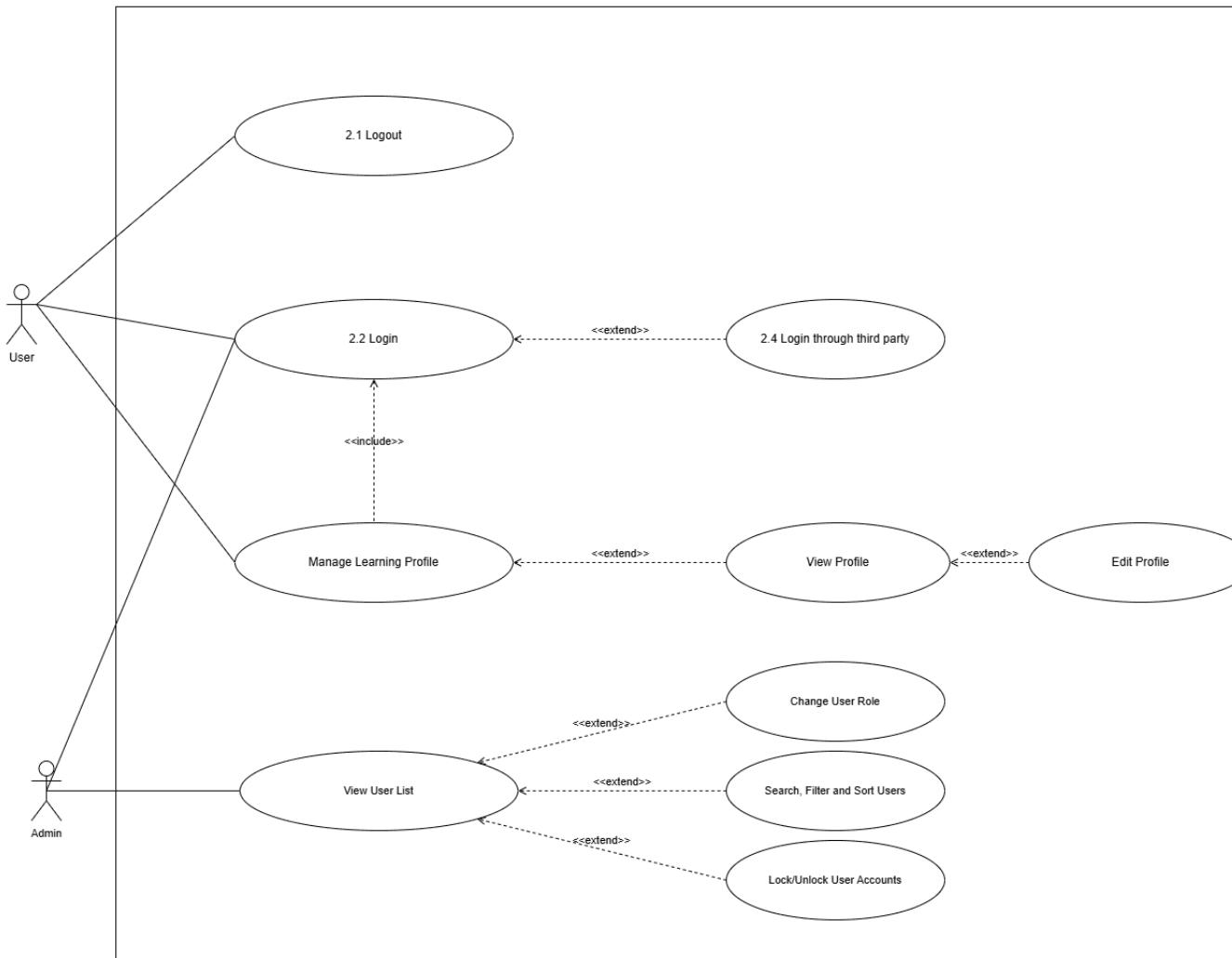


Figure 5.3: Detailed use case with id 1: Course Management

Use case **Quản lý khóa học** bao gồm ba chức năng chính: Tạo khóa học, Theo dõi kết quả học tập và Yêu cầu bài kiểm tra. Use case "**Chỉnh sửa khóa học**" được mở rộng từ "**Tạo khóa học**", cho thấy việc chỉnh sửa là một phần mở rộng tùy chọn của quá trình tạo khóa học. Use case diagram thể hiện các trách nhiệm chính của giáo viên trong việc quản lý khóa học trong hệ thống.

Ngoài ra, ở module này, Hệ thống sẽ có nhiệm vụ lưu trữ thông tin về khóa học mà giáo viên đã tạo ra.

### 5.1.2.2 User Management

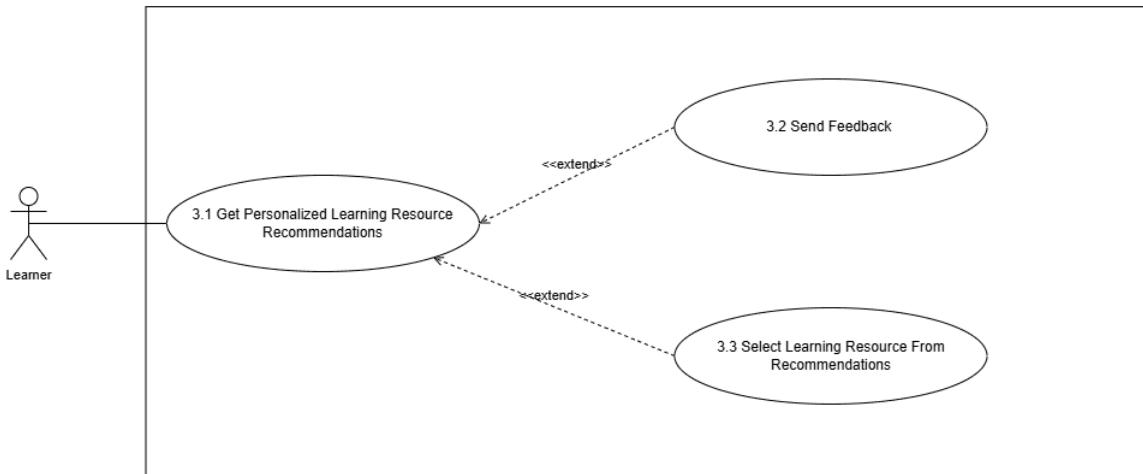


**Figure 5.4:** Detailed use case with id 2: User Management

Use case **Quản lý người dùng**, tập trung vào tác nhân Người dùng bao gồm các chức năng như Đăng nhập, Đăng xuất, Quản lý hồ sơ học tập và Xem hồ sơ. Use case diagram cho thấy "**Đăng nhập qua third-party**" mở rộng từ use case "**Đăng nhập**" cơ bản, cung cấp một phương thức đăng nhập thay thế. Ngoài ra, "**Xem hồ sơ**" được thể hiện như một phần mở rộng của "**Quản lý hồ sơ học tập**", cho thấy đây là một hành động liên quan nhưng riêng biệt.

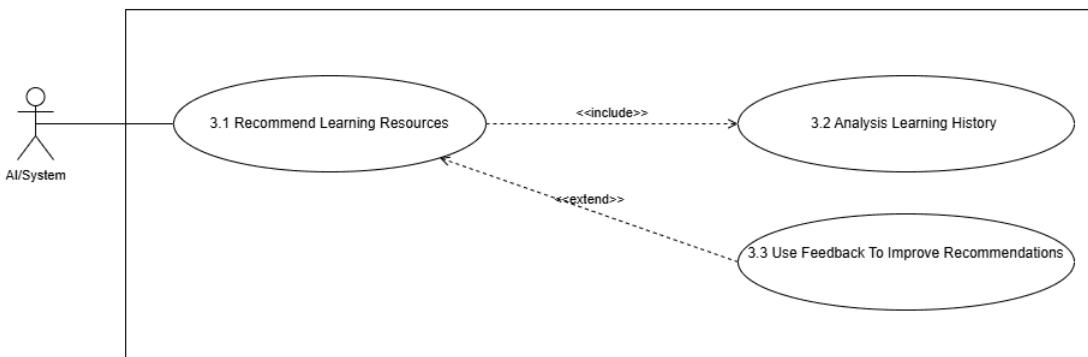
Quản trị viên có quyền đăng nhập bằng tài khoản của mình, từ đó "**Xem danh sách người dùng**" sử dụng hệ thống. Ngoài ra, quản trị viên có thể "**Tìm kiếm, Lọc và Sắp xếp theo thông tin người dùng**", "**Vô hiệu hóa tài khoản/Kích hoạt tài khoản**" và "**Thay đổi quyền truy cập của người dùng**". Ví dụ: Thay đổi từ người học sang giảng viên.

### 5.1.2.3 Learning Resource Recommendation System



**Figure 5.5:** Detailed use case with id 3: Adaptive Learning Path Management of Learner Actor

Tác nhân chính trong biểu đồ trên là người học (Learner), có thể tương tác với hệ thống để nhận các đề xuất tài nguyên học tập. Use case "**Get Personalized Learning Resource Recommendations**" cung cấp cho người học các tài nguyên học tập phù hợp với nhu cầu và lịch sử học tập của họ. Use case "**Send Feedback**" có mối quan hệ "<extend>" với "**Get Personalized Learning Resource Recommendations**", vì người học có thể gửi phản hồi về các tài nguyên đã được đề xuất để cải thiện hệ thống. Use case "**Select Learning Resource From Recommendations**" cũng mở rộng từ "**Get Personalized Learning Resource Recommendations**", cho phép người học tự chọn các tài nguyên mà họ thấy phù hợp với mục tiêu cá nhân.

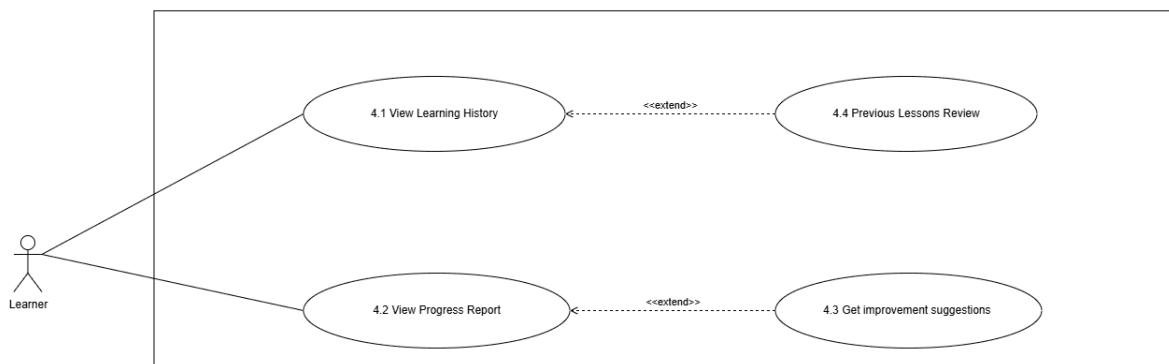


**Figure 5.6:** Detailed use case with id 3: Learning Resource Recommendation System of System Actor

Tác nhân chính là hệ thống AI, có nhiệm vụ chính là đề xuất tài nguyên học tập cá nhân hóa cho người dùng. Use case "**Recommend Learning Resources**" bao gồm hoạt động "**Analysis Learning History**" (phân tích lịch sử học tập), do quá trình phân tích lịch sử học tập là một phần không thể thiếu khi hệ thống tạo ra các đề xuất. Hệ thống có thể mở rộng thêm use case "**Use**

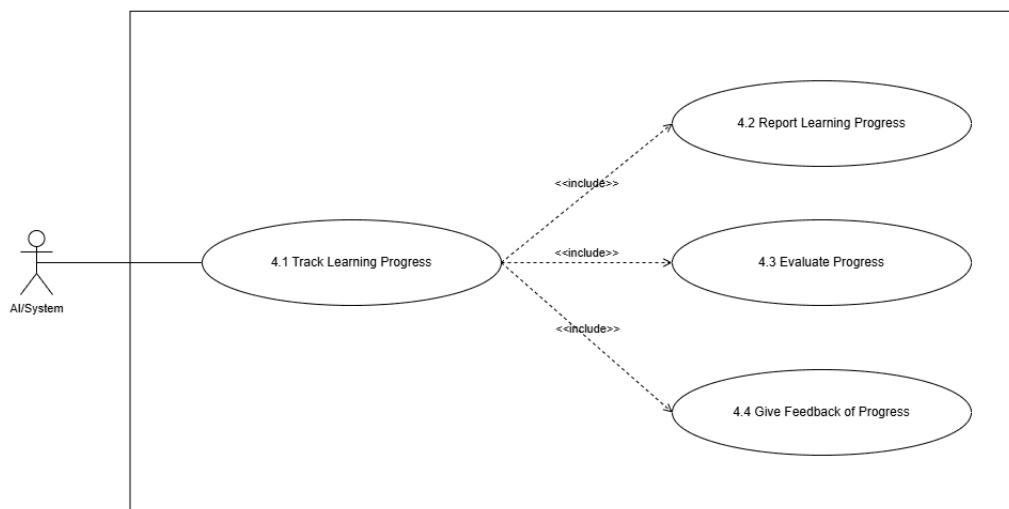
**Feedback To Improve Recommendations**" (sử dụng phản hồi để cải thiện đề xuất) dựa trên các phản hồi người dùng, giúp tinh chỉnh và nâng cao chất lượng đề xuất. Mỗi quan hệ "**<include>**" được sử dụng giữa "**Recommend Learning Resources**" và "**Analysis Learning History**" vì quá trình phân tích là cần thiết. Mỗi quan hệ "**<extend>**" giữa "**Recommend Learning Resources**" và "**Use Feedback To Improve Recommendations**" vì cải thiện đề xuất dựa trên phản hồi là một phần mở rộng có thể xảy ra tùy theo ngữ cảnh.

#### 5.1.2.4 Progress Tracking



**Figure 5.7:** Detailed use case with id 4: Progress Tracking of Learner Actor

Use case diagram trên thể hiện các chức năng **Theo dõi tiến độ dành cho Người học**. Có hai use case chính là "**Xem lịch sử học tập**" và "**Xem báo cáo tiến độ**". "**Xem lịch sử học tập**" mở rộng đến "**Xem lại các bài học trước đó**", cho phép người học ôn tập. "**Xem báo cáo tiến độ**" mở rộng đến "**Nhận đề xuất cải thiện**", cung cấp hướng dẫn cá nhân hóa, nhấn mạnh việc theo dõi và phản hồi liên tục để hỗ trợ quá trình học tập của người dùng.



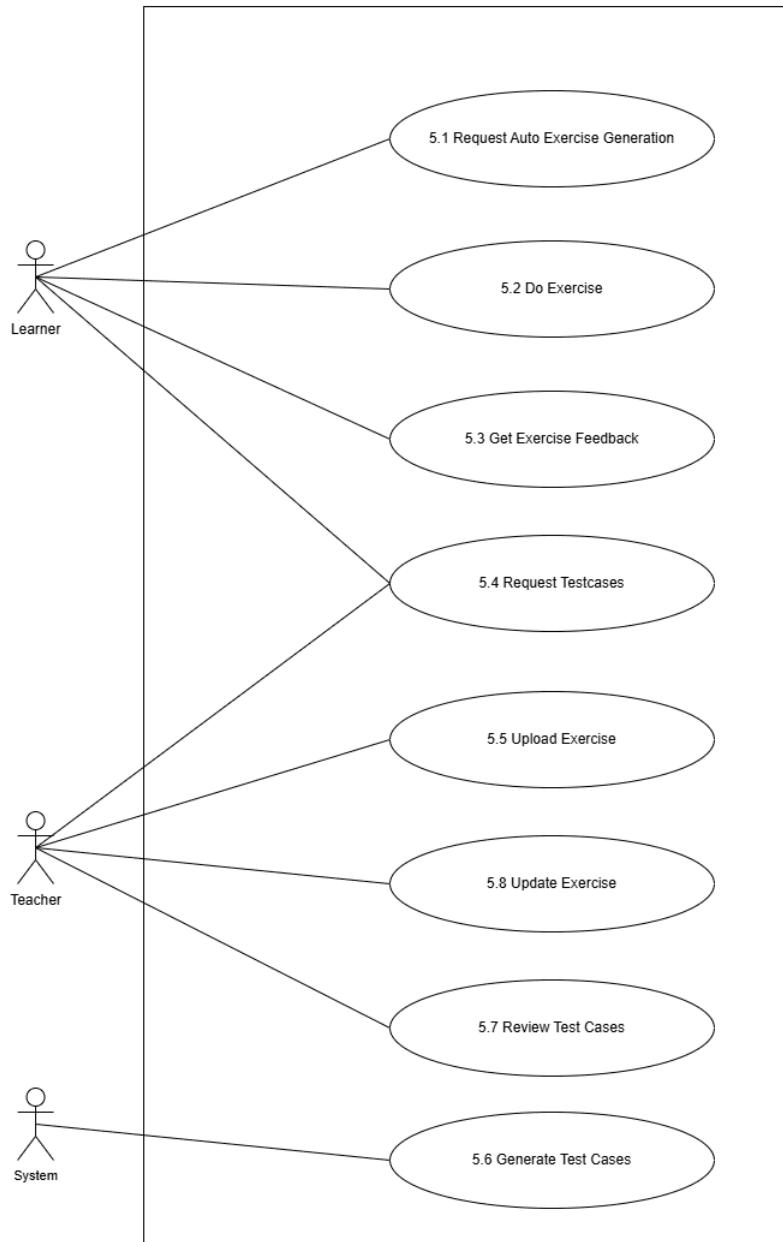
**Figure 5.8:** Detailed use case with id 4: Progress Tracking of System Actor

Sơ đồ use case của hệ thống **Exercise Generator** mô tả các tương tác giữa người dùng và



hệ thống. Người học có thể yêu cầu hệ thống tạo tự động các bài tập thông qua chức năng "**Request Auto Exercise Generation**", và sau đó thực hiện các bài tập đó để cải thiện kỹ năng của mình thông qua "**Do Exercise**". Bên cạnh đó, giảng viên có thể sử dụng chức năng "**Upload Exercise**" để tải lên các đề bài hoặc bài tập, từ đó hệ thống sẽ tự động xử lý và tạo ra các test case nhằm kiểm tra độ chính xác của bài giải thông qua "**Generate Test Cases**". Sau khi các test case được tạo, giảng viên có thể sử dụng chức năng "**Review Test Cases**" để xem xét và đánh giá các test case nhằm đảm bảo chúng phù hợp với nội dung bài tập và đáp ứng được yêu cầu giảng dạy. Các chức năng này giúp hệ thống tạo ra môi trường học tập linh hoạt, hỗ trợ hiệu quả cho người học và giảng viên.

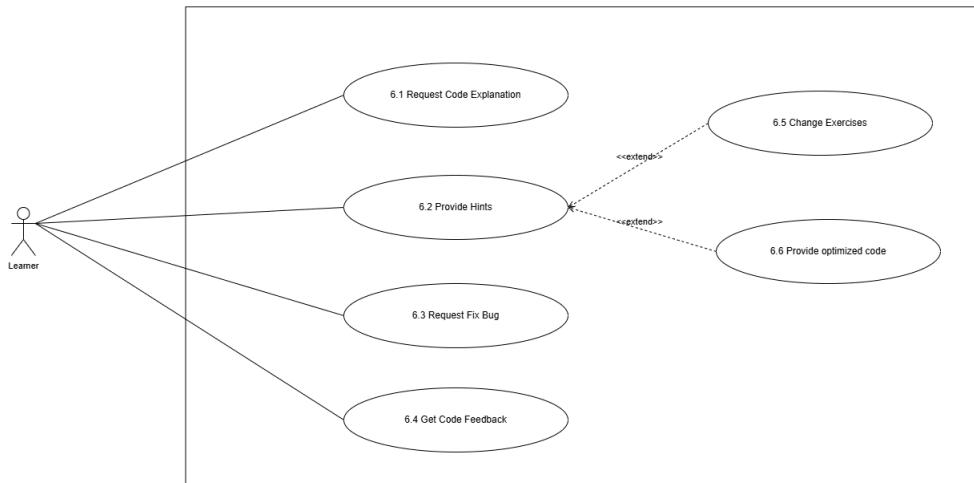
### 5.1.2.5 Exercise Generator



**Figure 5.9:** Detailed use case with id 5: Exercise Generator

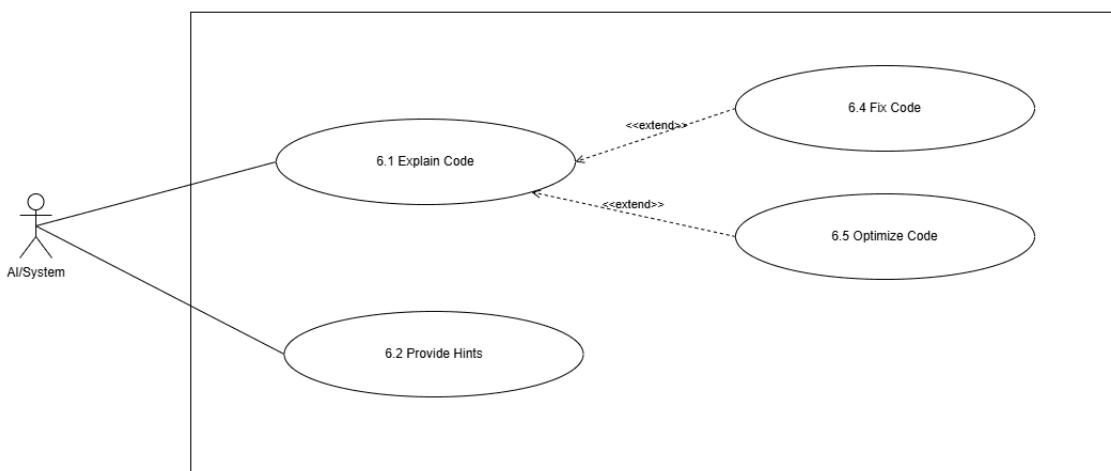
Use case diagram thể hiện các chức năng liên quan đến bài tập dành cho Người học. Có bốn use case chính: "*Yêu cầu tạo bài tập tự động*", "*Làm bài tập*", "*Nhận phản hồi về bài tập*", và "*Yêu cầu bài kiểm tra*". Người học có thể yêu cầu hệ thống tạo bài tập tự động phù hợp với trình độ của họ. Sau khi làm bài tập, người học có thể nhận phản hồi để hiểu rõ hơn về kết quả của mình. Khả năng yêu cầu bài kiểm tra cho phép người học tự đánh giá kiến thức của mình khi cần.

### 5.1.2.6 AI Tutor



**Figure 5.10:** Detailed use case with id 6: AI Tutor of Learner Actor

Use case diagram thể hiện các chức năng của Gia sư AI dành cho Người học. Có bốn use case chính: "*Yêu cầu giải thích mã*", "*Yêu cầu gợi ý*", "*Yêu cầu sửa lỗi*", và "*Nhận phản hồi về mã*". Người học có thể yêu cầu giải thích về mã nguồn để hiểu rõ hơn về cấu trúc và logic của code. Chức năng yêu cầu gợi ý có hai chức năng quan hệ "*extend*" là "*Hỗ trợ chuyển đổi bài tập*" dựa trên nhu cầu, học lực của người dùng và "*Cung cấp mã tối ưu hóa*" giúp người học vượt qua các khó khăn trong quá trình học lập trình cũng như cải thiện tối đa kỹ thuật code của người dùng. Việc nhận phản hồi về mã giúp người học cải thiện kỹ năng lập trình của mình một cách liên tục.



**Figure 5.11:** Detailed use case with id 6: AI Tutor of System Actor

Use case diagram trên mô tả **hệ thống AI Tutor** với vai trò hỗ trợ người dùng trong việc xử lý và hiểu mã lập trình. Hệ thống có khả năng *giải thích mã lệnh* (*Use case 6.1*) và *cung cấp gợi ý khi người dùng gặp khó khăn* (*Use case 6.2*). Từ chức năng giải thích mã, hệ thống có



thể mở rộng thêm các chức năng khác như *sửa lỗi* (*Use case 6.4*) và *tối ưu mã lệnh* (*Use case 6.5*). Hệ thống nhằm mục tiêu hỗ trợ người dùng học lập trình hiệu quả hơn thông qua các công cụ hỗ trợ trực tiếp từ AI.



## 5.2 Thiết kế

### 5.2.1 Kiến trúc hệ thống

#### 5.2.1.1 Phân tích kiến trúc hệ thống phổ biến hiện nay

Hiện nay, có nhiều kiến trúc phần mềm khác nhau, mỗi loại có ưu và nhược điểm riêng, phù hợp với các loại dự án và yêu cầu cụ thể. Dưới đây là một số kiến trúc phần mềm phổ biến và các điểm chính của chúng:

##### 1. Kiến trúc Monolithic

- **Định nghĩa:** Kiến trúc monolithic là một mô hình truyền thống trong đó toàn bộ ứng dụng được xây dựng như một đơn vị duy nhất, tích hợp chặt chẽ.
- **Điểm chính:**
  - Đơn giản để phát triển, triển khai và kiểm thử
  - Hiệu suất tốt do không có overhead của network calls
  - Khó mở rộng và bảo trì khi ứng dụng phức tạp
  - Khó áp dụng công nghệ mới cho từng phần của ứng dụng
- **Ví dụ:** Một ứng dụng web PHP truyền thống, nơi tất cả chức năng (UI, business logic, data access) đều nằm trong cùng một codebase.

##### 2. Kiến trúc Microservices

- **Định nghĩa:** Kiến trúc microservices chia nhỏ ứng dụng thành các dịch vụ độc lập, mỗi dịch vụ chịu trách nhiệm cho một chức năng cụ thể và có thể được phát triển, triển khai độc lập.
- **Điểm chính:**
  - Dễ dàng mở rộng và bảo trì từng service riêng biệt
  - Cho phép sử dụng công nghệ phù hợp nhất cho mỗi service
  - Tăng cường khả năng chịu lỗi của hệ thống
  - Phức tạp hơn trong việc quản lý và điều phối các service
- **Ví dụ:** Netflix sử dụng kiến trúc microservices, với các service riêng biệt cho việc xử lý video, quản lý người dùng, đề xuất nội dung, v.v.

##### 3. Kiến trúc Client-Server (Multitier)

- **Định nghĩa:** Kiến trúc client-server chia ứng dụng thành các tầng (tiers) riêng biệt, thường bao gồm tầng trình bày (presentation), tầng logic nghiệp vụ (business logic), và tầng dữ liệu (data).



• **Điểm chính:**

- Phân tách rõ ràng giữa các tầng chức năng
  - Dễ dàng mở rộng từng tầng độc lập
  - Cải thiện bảo mật bằng cách kiểm soát quyền truy cập giữa các tầng
  - Có thể trở nên phức tạp khi số lượng tầng tăng lên
- **Ví dụ:** Một ứng dụng web ba tầng điển hình bao gồm frontend (React), backend API (Node.js), và cơ sở dữ liệu (PostgreSQL).

#### 4. Kiến trúc Event-Driven

- **Định nghĩa:** Kiến trúc event-driven là một mô hình trong đó việc tạo ra, phát hiện, tiêu thụ và phản ứng với các sự kiện là cốt lõi của hệ thống.
  - **Điểm chính:**
    - Tách rời các thành phần hệ thống, giảm sự phụ thuộc trực tiếp
    - Khả năng mở rộng và linh hoạt cao
    - Dễ dàng thêm chức năng mới mà không ảnh hưởng đến các thành phần hiện có
    - Có thể phức tạp trong việc theo dõi luồng xử lý và debug
- **Ví dụ:** Hệ thống thương mại điện tử, khi một đơn hàng được đặt, một sự kiện "OrderPlaced" được phát ra, kích hoạt các quy trình xử lý khác như cập nhật kho, gửi email xác nhận, v.v.

#### 5. Kiến trúc Serverless

- **Định nghĩa:** Kiến trúc serverless cho phép phát triển và chạy ứng dụng mà không cần quản lý trực tiếp máy chủ, tập trung vào việc viết code cho các chức năng (functions) cụ thể.
  - **Điểm chính:**
    - Giảm chi phí vận hành và bảo trì hạ tầng
    - Tự động mở rộng theo nhu cầu sử dụng
    - Chỉ trả tiền cho tài nguyên thực sự sử dụng
    - Có thể gặp vấn đề về cold start và vendor lock-in
- **Ví dụ:** Một ứng dụng xử lý ảnh sử dụng AWS Lambda để thực hiện các tác vụ như resize, filter, và lưu trữ ảnh khi người dùng tải lên.

##### 5.2.1.2 Lựa chọn kiến trúc hệ thống

Dựa trên yêu cầu của hệ thống, nhóm sử dụng kiến trúc kết hợp giữa Microservices và Multitier Client-Server. Cụ thể như sau:



1. Presentation Tier (Client Tier):

- Frontend: Vue.js application
- Giao diện người dùng cho sinh viên và giảng viên
- Tương tác với backend thông qua RESTful API

2. Application Tier:

(a) API Gateway:

- Nginx hoặc API Gateway mã nguồn mở
- Xử lý authentication, authorization và định tuyến requests đến các microservices

(b) Microservices (tất cả được xây dựng bằng Fast API):

- User Service:
  - Quản lý thông tin người dùng
  - Xác thực và phân quyền
- Course Management Service:
  - CRUD operations cho khóa học, module, bài học
  - Quản lý nội dung học tập
- Recommendation Learning Service:
  - Đề xuất learning items dựa trên dữ liệu người dùng
- AI Tutor Service:
  - Tích hợp Langchain để xây dựng gia sư AI
  - Xử lý các yêu cầu liên quan đến hỗ trợ lập trình và giải thích mã nguồn
- Exercise Generation Service:
  - Sử dụng Langchain để tạo bài tập và test cases tự động
  - Đánh giá bài làm của người học
- Progress Tracking Service:
  - Theo dõi và ghi nhận tiến độ học tập
  - Tạo báo cáo học tập

3. Data Tier:

(a) Persistence Layer:

- PostgreSQL: Lưu trữ dữ liệu có cấu trúc (thông tin người dùng, khóa học, bài học)



- Neo4j: Lưu trữ và quản lý mối quan hệ giữa các learning items, hỗ trợ tạo lộ trình học tập
- AWS S3: Lưu trữ tài liệu học tập, video bài giảng

(b) Caching Layer:

- Redis: Cải thiện hiệu suất và giảm tải cho database

**Phân tích nguyên nhân:**

- Phân chia trách nhiệm rõ ràng: Mỗi dịch vụ có một nhiệm vụ cụ thể, dễ bảo trì và phát triển. Ví dụ, FastAPI quản lý API, Langchain xử lý AI, PostgreSQL và Neo4j quản lý dữ liệu.
- Dễ mở rộng: Mỗi dịch vụ có thể mở rộng riêng lẻ khi tải tăng lên, như tăng số lượng instances của FastAPI khi số người dùng tăng.
- Triển khai linh hoạt: Thay đổi hoặc thêm công nghệ mới dễ dàng mà không ảnh hưởng đến toàn hệ thống.

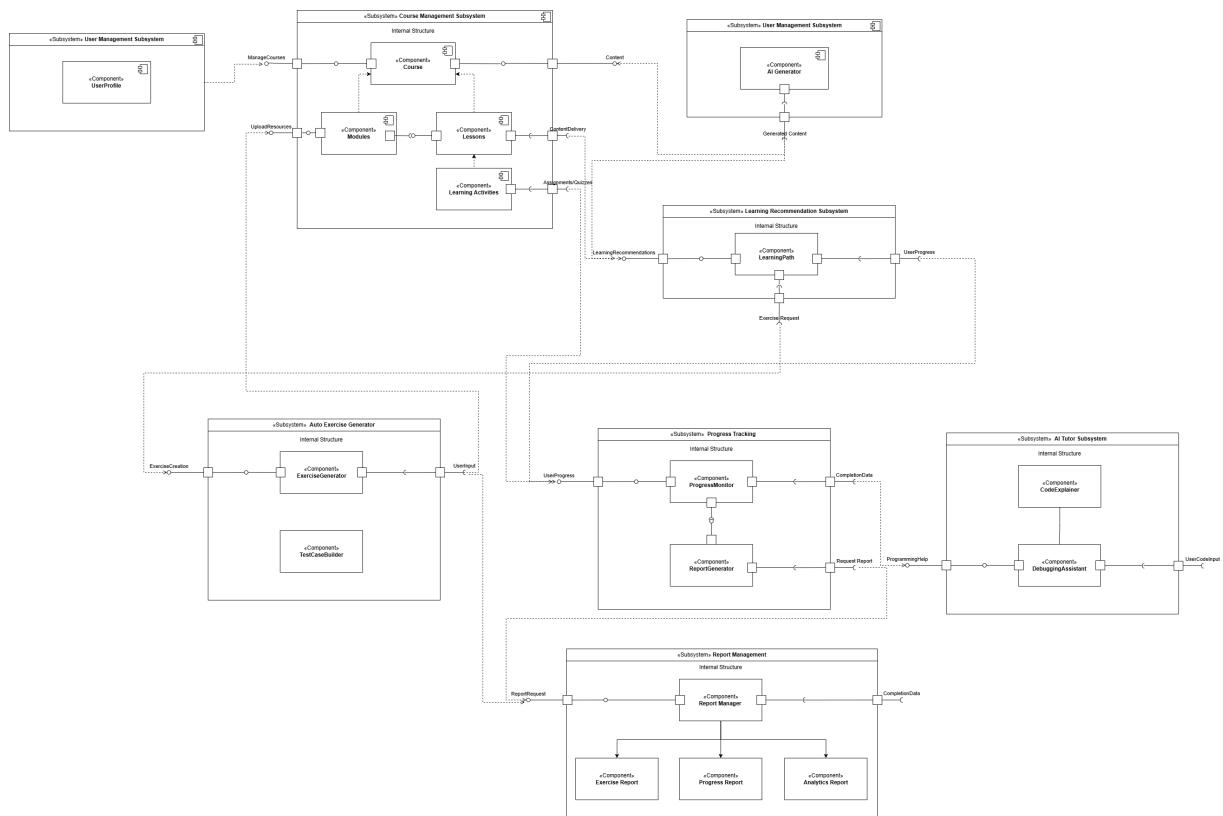
**Không dùng:**

- Monolithic vì khó mở rộng, bảo trì, và tích hợp công nghệ mới.
- Event-driven vì phức tạp khi quản lý trạng thái trực tiếp và đồng bộ.
- Serverless do phụ thuộc nhà cung cấp, chi phí cao khi tải tăng, và hạn chế tùy chỉnh hiệu suất.

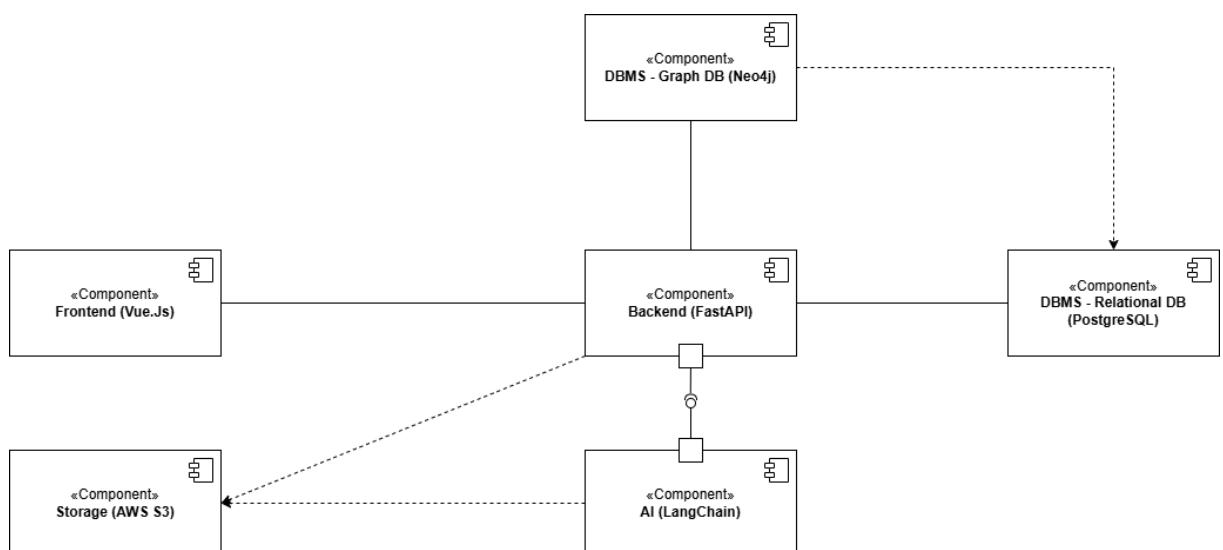
Như vậy, ta có luồng hoạt động của hệ thống như sau:

- Người dùng tương tác với frontend Vue.js (Presentation Tier).
- Requests được gửi đến API Gateway trong Application Tier.
- API Gateway xác thực và định tuyến requests đến microservice phù hợp.
- Microservices trong Application Tier xử lý business logic:
  - Gọi đến Persistence Layer trong Data Tier để lưu trữ hoặc truy xuất dữ liệu.
  - Sử dụng Caching Layer để tối ưu hiệu suất truy vấn.
- Kết quả được trả về cho frontend thông qua API Gateway.

**Tóm lại**, kiến trúc kết hợp ưu điểm của Microservices và Multitier tận dụng tối đa các công nghệ đã được chọn để xây dựng một hệ thống học tập trực tuyến thông minh, có khả năng mở rộng cao và linh hoạt. Việc phân chia rõ ràng giữa các tiers và microservices giúp dễ dàng phát triển, bảo trì và mở rộng hệ thống trong tương lai.



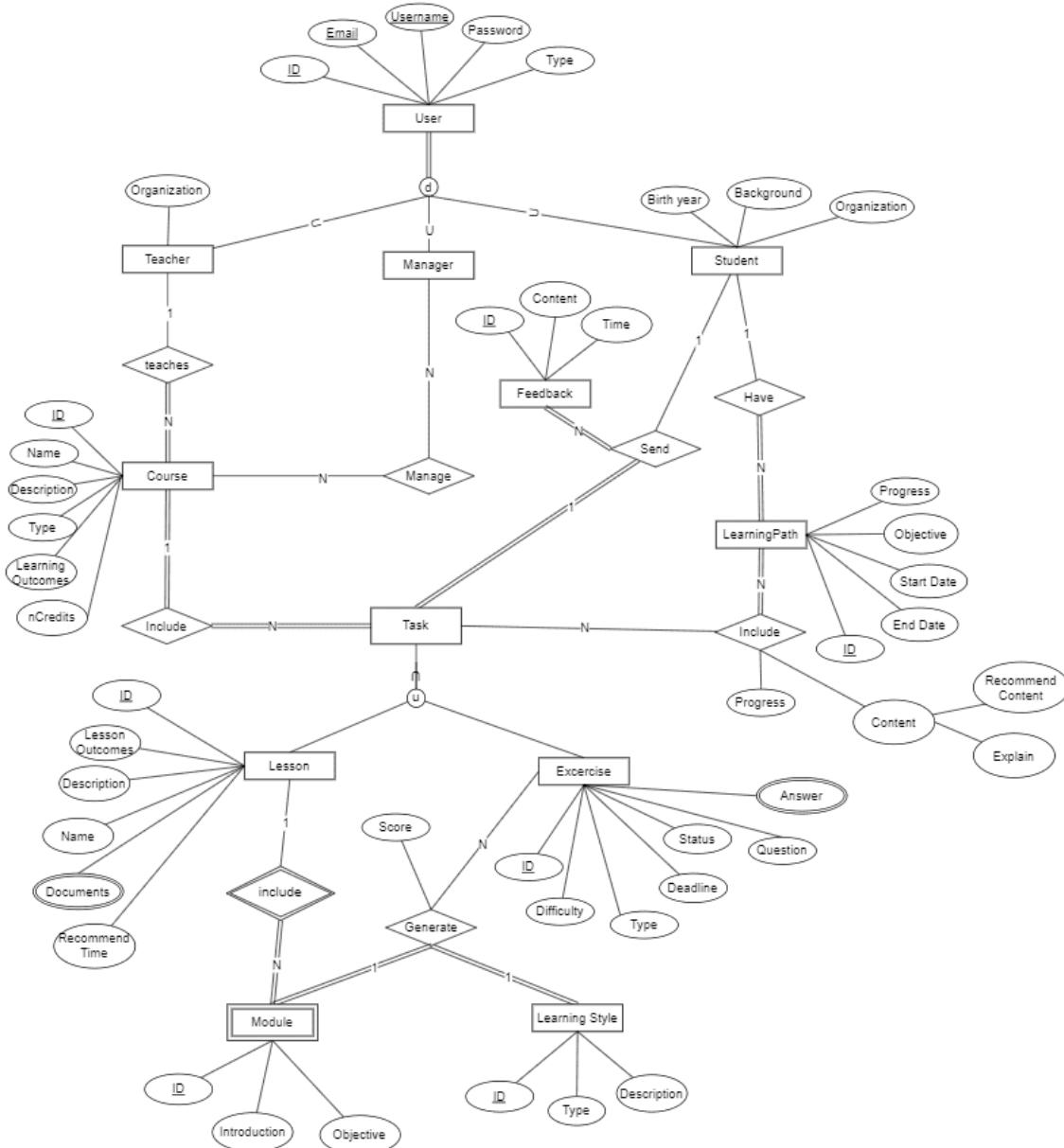
**Figure 5.12: Component-based Diagram cho toàn bộ hệ thống**



**Figure 5.13: Component-based Diagram cho những công nghệ được chọn**

## 5.3 Database

### 5.3.1 EER diagram



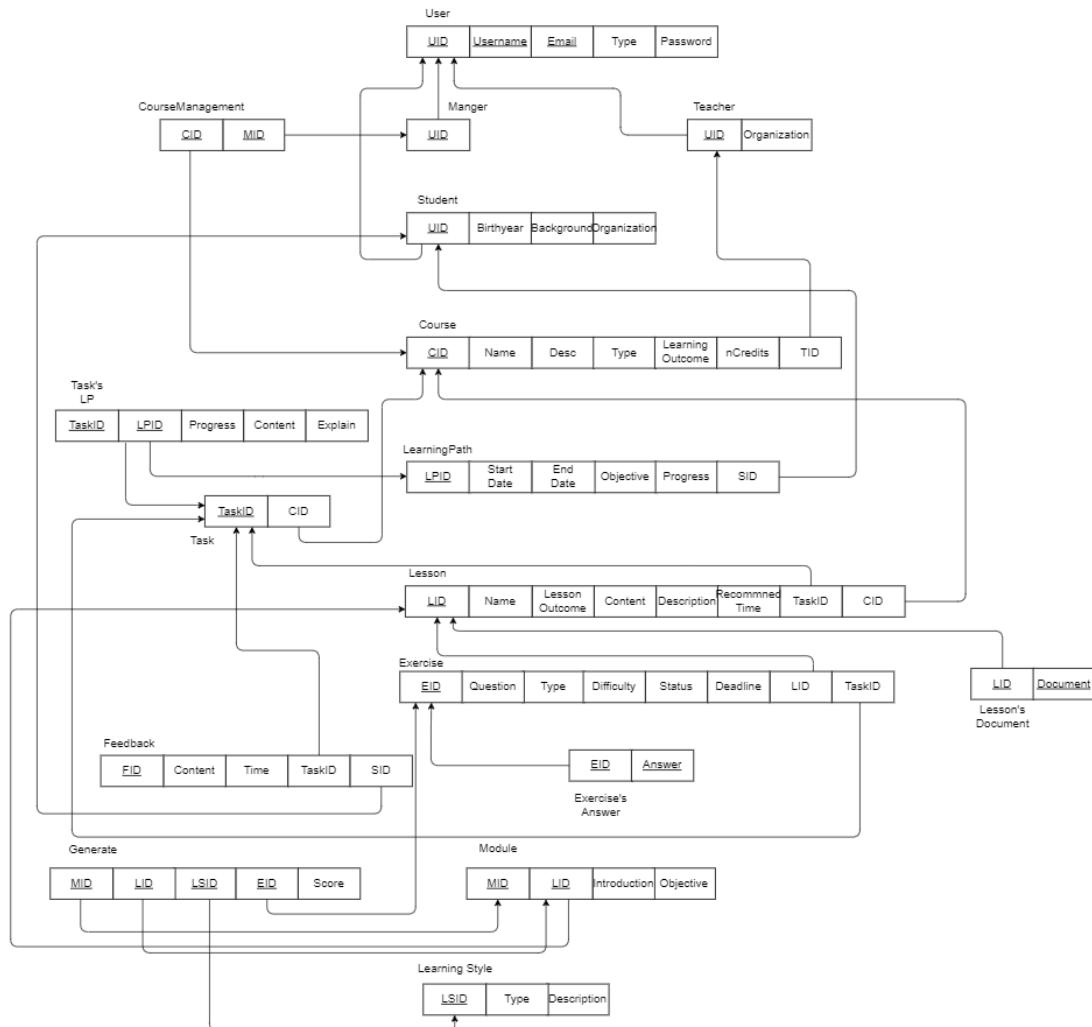
**Figure 5.14: Sơ đồ EER của hệ thống**

1. User có ba loại con: Teacher, Manager, và Student. Mỗi loại người dùng có các mối quan hệ khác nhau với các thực thể khác trong hệ thống:
  - Student có các mối quan hệ sau:
    - Include một hoặc nhiều Activity.
    - Send nhiều Feedback cho hệ thống.



- Submit nhiều Exercise để thực hành.
  - Join nhiều Course để tham gia các khóa học.
  - Bookmark nhiều Lesson để lưu trữ các bài học quan trọng.
- Teacher có thể Upload nhiều Course vào hệ thống để giảng dạy.
  - Manager có thể Manage nhiều Course trong hệ thống, tức là có quyền quản lý các khóa học.
2. Course bao gồm nhiều Lesson, và mỗi khóa học có thể Include nhiều bài học khác nhau. Mỗi học viên (Student) có thể tham gia một khóa học và cho mỗi khóa học, họ sẽ có một LearningPath riêng biệt. Mỗi quan hệ này thể hiện rằng mỗi học viên có một lộ trình học tập cho mỗi khóa học mà họ tham gia.
3. Lesson có thể Specialize thành Recommend Lesson, tức là các bài học có thể bao gồm các đề xuất học tập cho học viên. Mỗi bài học cũng có thể Include nhiều Module.
4. LearningPath của mỗi học viên sẽ Have nhiều Recommend Lesson. Các lộ trình học tập này sẽ dẫn dắt học viên thông qua các bài học được đề xuất trong suốt quá trình học.
5. Recommend Lesson có thể Include nhiều Module, tức là mỗi bài học được đề xuất có thể bao gồm nhiều mô-đun khác nhau để hỗ trợ học viên học tập một cách chi tiết hơn.
6. Module có thể Generate nhiều Quiz để đánh giá kết quả học tập của học viên. Mỗi mô-đun cũng có thể Generate một tài liệu (Document) duy nhất liên quan đến nội dung học.
7. Quiz chứa các câu hỏi và bài kiểm tra học viên, có thể bao gồm các phản hồi từ LLM Response, cung cấp thông tin về tên, độ khó, câu hỏi, câu trả lời, các lựa chọn, và hình ảnh liên quan đến bài kiểm tra.
8. Document có thể được tạo ra từ các mô-đun và cũng có phản hồi từ LLM Response, giúp cung cấp thêm tài liệu học tập cho học viên.
9. Activity bao gồm các hành động của người dùng trong hệ thống, và mỗi hoạt động có thể được Include bởi một hoặc nhiều User, đặc biệt là học viên.
10. Feedback là thông tin phản hồi từ học viên gửi đến hệ thống và có thể được Send từ nhiều User, đặc biệt là từ học viên để cải thiện chất lượng giảng dạy và học tập.

### 5.3.2 Lược đồ CSDL quan hệ ánh xạ

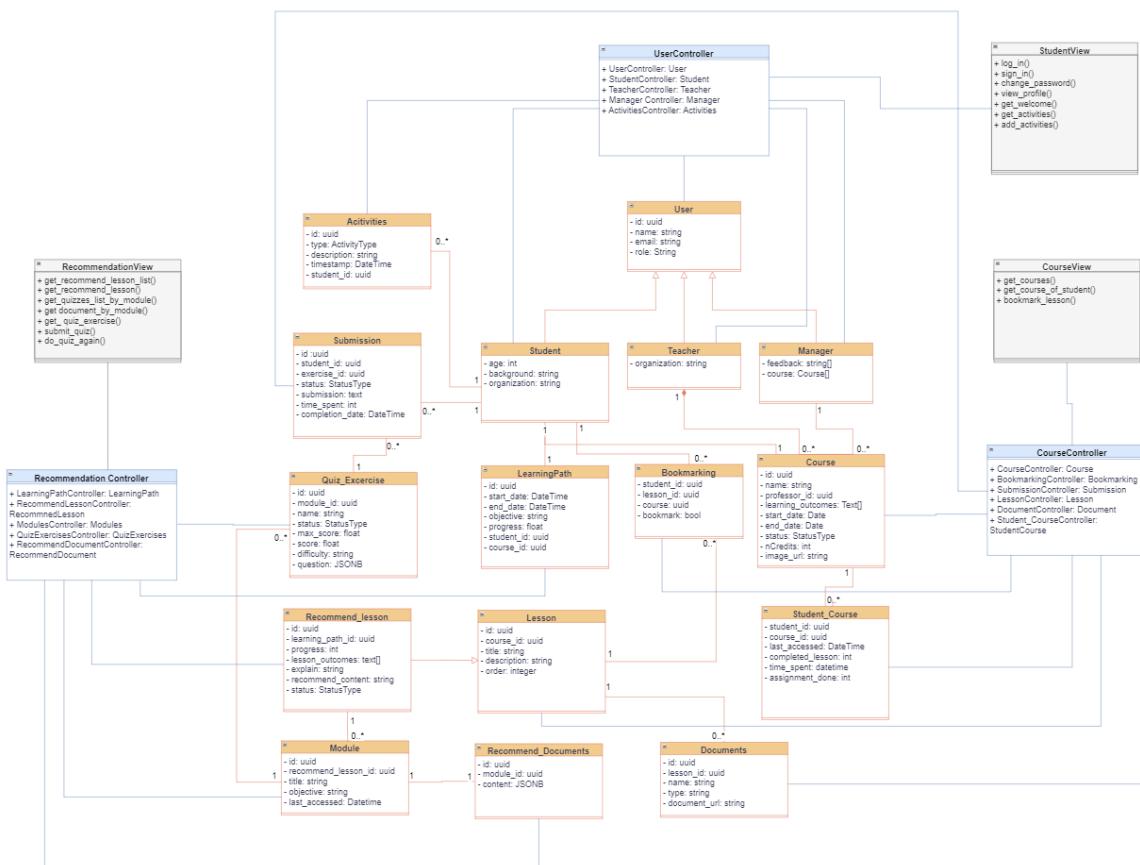


**Figure 5.15: Lược đồ CSDL quan hệ ánh xạ của hệ thống**

Lược đồ cơ sở dữ liệu ánh xạ biểu diễn các mối quan hệ giữa các thực thể chính trong hệ thống quản lý học tập thông minh. Các thực thể như User (bao gồm Student, Teacher, và Manager), Course, LearningPath, Lesson, Module, và Quiz được tổ chức thành các bảng cơ sở dữ liệu, thể hiện mối quan hệ và tương tác của chúng.

Lược đồ cơ sở dữ liệu này không chỉ giúp tổ chức dữ liệu theo một cấu trúc rõ ràng mà còn giúp xác định các ràng buộc và quan hệ giữa các thực thể khác nhau trong hệ thống.

### 5.3.3 Class diagram



**Figure 5.16: Class diagram hiện thực của hệ thống**

Class diagram được thiết kế theo mô hình MVC mô tả một hệ thống học tập thông minh với ba vai trò chính là sinh viên, giáo viên và quản trị viên, tất cả đều kế thừa từ lớp người dùng (User). Mỗi người dùng có các chức năng cơ bản như đăng nhập, đăng xuất và cập nhật hồ sơ. Sinh viên chọn theo học khóa học của giảng viên, thực hiện bài tập, và theo dõi tiến độ học tập của mình. Lộ trình học tập bao gồm bài tập được cá nhân hóa cho từng sinh viên. Giáo viên chịu trách nhiệm tạo và quản lý các khóa học, bao gồm việc thêm bài tập, theo dõi kết quả học tập của sinh viên, và tạo ra các bộ câu hỏi kiểm tra. Quản trị viên có quyền quản lý toàn bộ hệ thống, bao gồm khóa học và người dùng, cũng như giải quyết các thắc mắc từ người dùng. AI được sử dụng để hỗ trợ quá trình học tập thông qua việc tự động tạo lộ trình học tập và bài tập. AI dựa vào dữ liệu từ các tài liệu học tập, câu hỏi và đáp án có sẵn, cũng như các bài tập lập trình để cá nhân hóa trải nghiệm học cho người dùng.

### 5.3.4 Sitemap

#### 5.3.4.1 Sinh viên

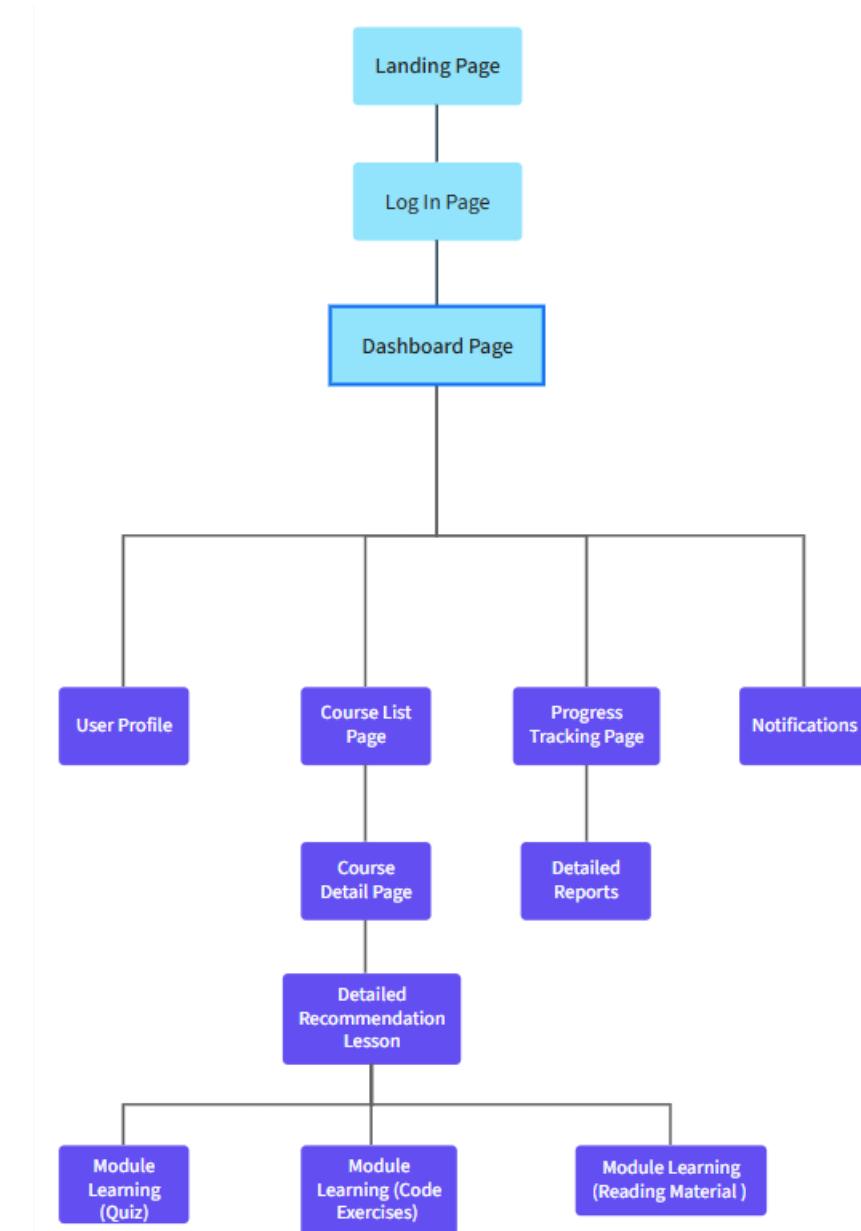


Figure 5.17: Sitemap cho đối tượng Sinh viên

Sơ đồ trang web này thể hiện cấu trúc tổng thể của ứng dụng web, hiển thị các trang chính và mối quan hệ giữa chúng, đặc biệt là đối với đối tượng người dùng **Sinh viên**. Ở phía trên, có "Landing Page", "Log In Page" và "Dashboard Page". Đây là các trang điều hướng chính cho người dùng.

Bên dưới, có một số phần khác:



- **User Profile:** Phần này chứa các trang liên quan đến tài khoản, thông tin cá nhân và cài đặt của sinh viên.
- **Course List Page:** Trang này hiển thị danh sách các khóa học mà sinh viên có thể truy cập.
- **Progress Tracking Page:** Trang này cho phép sinh viên theo dõi tiến độ và hiệu suất của họ trong các khóa học.
- **Notifications:** Phần này bao gồm trang chứa các thông báo cho sinh viên các cập nhật, tin nhắn hoặc cảnh báo quan trọng.
- **Course Detail Page:** Trang này cung cấp thông tin chi tiết về một khóa học cụ thể, như nội dung, mục tiêu và các tài nguyên khác.
- **Detailed Recommendation Lesson:** Trang này cung cấp các khóa học đề xuất học tập dựa trên tiến độ và mong muốn của người dùng.
- **Module Learning (Quiz), Module Learning (Code Exercises), and Module Learning (Reading Material):** Những phần này đại diện cho nội dung module của các bài học được đề xuất và chia thành 3 dạng bài tập như bài kiểm tra dạng quiz (multiple choices), bài tập lập trình thực hành và tài liệu đọc.

Nhìn chung, sơ đồ trang web này hiển thị các trang của hệ thống của đối tượng người dùng **Sinh viên** và cung cấp một cái nhìn tổng quan về cách các trang và chức năng liên quan được tổ chức và tương tác với nhau.

#### 5.3.4.2 Giảng viên

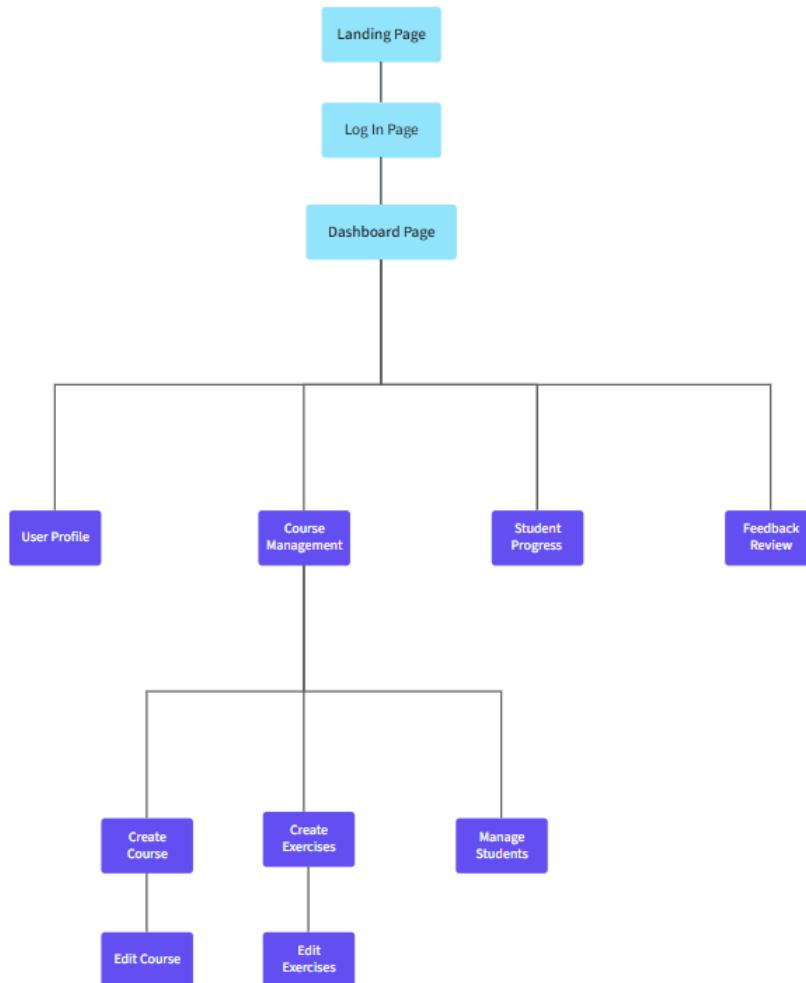


Figure 5.18: Sitemap cho đối tượng Giảng viên

Sơ đồ trang web này đại diện cho cấu trúc của ứng dụng web, tập trung vào các chức năng và trang dành riêng cho giảng viên.

Ở phía trên, vẫn có các trang điều hướng chính như "Landing Page", "Log In Page" và "Dashboard Page", giúp giảng viên dễ dàng truy cập vào các phần quan trọng của hệ thống.

Bên dưới, sơ đồ được chia thành bốn phần chính:

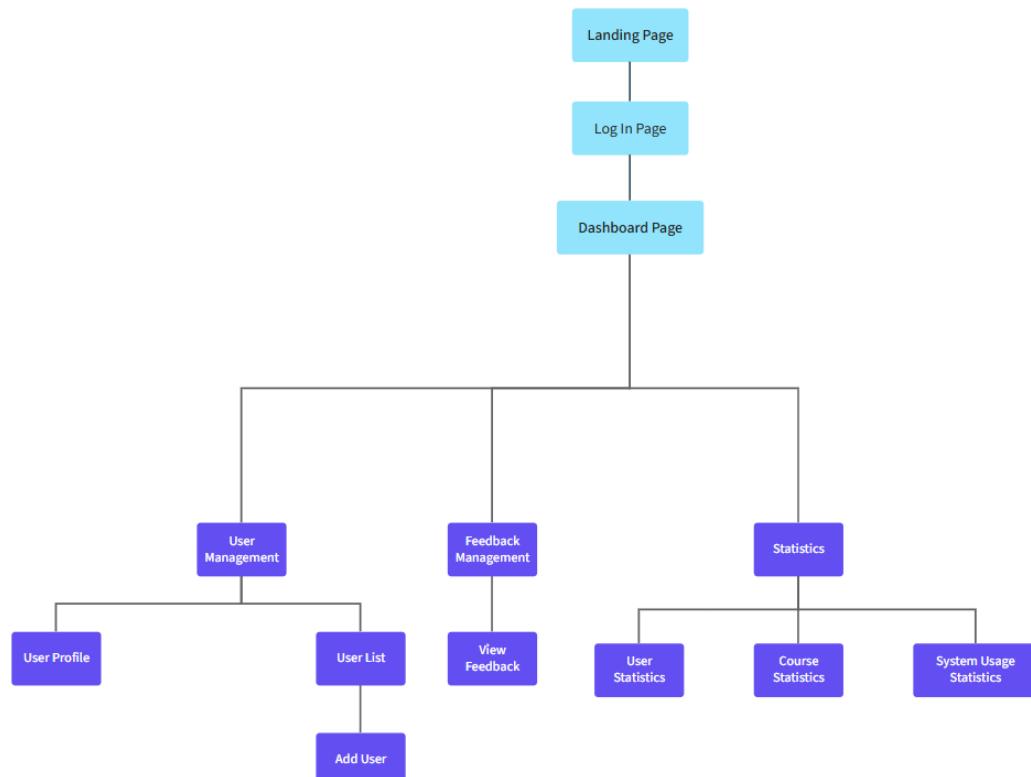
- **User Profile (Hồ sơ người dùng):** Phần này chứa các trang liên quan đến thông tin cá nhân, tài khoản và cài đặt của giảng viên. Giảng viên có thể cập nhật thông tin cá nhân và quản lý các tùy chọn hệ thống của mình.
- **Course Management (Quản lý khóa học):** Đây là phần quan trọng nhất với giảng viên, bao gồm các trang để tạo, chỉnh sửa, và quản lý các khóa học mà họ giảng dạy. Giảng viên có thể thêm các bài giảng, tài liệu học, và bài kiểm tra cho sinh viên.

- **Student Progress (Tiến độ học sinh):** Phần này giúp giảng viên theo dõi và giám sát tiến độ học tập của sinh viên trong các khóa học mà họ giảng dạy. Giảng viên có thể xem thông tin về kết quả học tập của sinh viên, bao gồm điểm kiểm tra, bài tập và các hoạt động học tập khác.
- **Feedback Management (Quản lý phản hồi):** Phần này cho phép giảng viên thu thập, quản lý và phân tích phản hồi từ sinh viên về khóa học, bài giảng, hoặc các hoạt động học tập. Đây là công cụ hữu ích để giảng viên đánh giá mức độ hiệu quả của phương pháp giảng dạy và cải thiện chất lượng khóa học.

Các trang con trong mỗi phần cung cấp các chức năng cụ thể hơn, chẳng hạn như chỉnh sửa nội dung khóa học, xem bảng điểm và thông tin phản hồi từ sinh viên.

**Tổng quan:** Sơ đồ trang web này cung cấp cái nhìn tổng quát về các chức năng chính mà giảng viên cần để quản lý khóa học, theo dõi tiến độ sinh viên và thu thập phản hồi. Cấu trúc đơn giản, dễ sử dụng, giúp giảng viên dễ dàng thực hiện các công việc quản lý và giảng dạy hiệu quả.

#### 5.3.4.3 Manager



**Figure 5.19:** Sitemap cho đối tượng Quản trị viên

Sơ đồ trang web này cung cấp cái nhìn tổng quát về các chức năng chính mà Admin cần để quản lý các chức năng người dùng.



Ở phía trên, các trang điều hướng chính vẫn là "Landing Page", "Log In Page", và "Dashboard Page".

Các phần chính trong sơ đồ trang web này là:

- **User Management:** Phần này bao gồm các trang để quản lý hồ sơ cá nhân, cài đặt và thông tin tài khoản của người dùng.
- **Feedback Management:** Phần này tập trung vào các trang để cung cấp phản hồi, đánh giá và nhận xét liên quan đến trải nghiệm học tập của người dùng.
- **Statistics:** Phần này cung cấp các trang hiển thị các số liệu thống kê, chỉ số và dữ liệu liên quan đến tiến độ, hiệu suất và mức độ sử dụng hệ thống của người dùng.

Các trang cấp thấp hơn trong sơ đồ trang web này hướng đến học sinh, chẳng hạn như các trang để xem chi tiết khóa học, truy cập tài liệu học tập và theo dõi tiến độ cá nhân.

Sơ đồ trang web này gợi ý về một ứng dụng web được thiết kế để cung cấp cho học sinh một trải nghiệm học tập cá nhân hóa, với sự tập trung vào quản lý người dùng, phản hồi và các thông tin chi tiết dựa trên dữ liệu để hỗ trợ hành trình giáo dục của họ.

## Tổng kết

Nhìn chung, ba sơ đồ trang web này thể hiện các góc nhìn và ưu tiên khác nhau trong thiết kế và chức năng của ứng dụng web, đáp ứng nhu cầu của các loại người dùng khác nhau, chẳng hạn như học sinh, giảng viên và quản trị viên.

## 5.4 Thiết kế prompt để xây dựng cơ sở dữ liệu đồ thị và hệ thống đề xuất lộ trình học cho từng khóa học

Trong hệ thống học tập này, LLM (Language Model) được sử dụng để hỗ trợ việc trích xuất thông tin về người học và tài nguyên học tập, từ đó giúp đề xuất các tài nguyên học tập phù hợp với mục tiêu học tập của người dùng. Quy trình này được thực hiện thông qua một chuỗi các bước được mô tả chi tiết dưới đây.

## 5.5 Thiết kế prompt để xây dựng đồ thị tri thức từ nguồn học liệu (bài giảng, module,...)

Việc xây dựng đồ thị từ nguồn học liệu sẽ được thực hiện bởi mô hình ngôn ngữ lớn, và dữ liệu sẽ được thêm vào cơ sở dữ liệu đồ thị bằng kỹ thuật *function calling*:

Context: You are an assistant helping teachers and instructors process their learning resources.



Your role is to analyze and extract data from various learning materials (such as courses, lessons, learning modules, assignments, exercises, etc.) and convert it into a format suitable for insertion into a graph database. This database will support a personalized learning resource recommendation system tailored for students.

Your main task is to use the tool ‘AddLearningResource‘ to categorize each learning resource and identify relevant concepts. For each learning resource, you must identify and list at least 5 related concepts that represent the key ideas or topics covered.

Instructions:

1. Carefully examine the provided JSON data and understand the main topics of the resource.
2. Use ‘add\_learning\_resource‘ to log the resource details and associated concepts.
3. Ensure each resource has at least 5 related concepts to maximize the utility of the recommendation system.
4. If the input data is a JSON array, please call AddLearningResource function for each learning resource in the array. Output your findings in a structured format.

Câu prompt trên sử dụng function calling để mô hình có thể tự động thực hiện các tác vụ sau khi phân tích tài nguyên học tập, trong đó:

- Mô hình nhận diện các khái niệm chính từ nội dung của tài liệu học tập, chẳng hạn như "Networking", "Machine Learning", và các chủ đề liên quan. Các khái niệm này giúp xây dựng một cấu trúc đồ thị tri thức, từ đó hỗ trợ hệ thống gợi ý tài nguyên học tập cá nhân hóa cho người học.
- Sau khi trích xuất các khái niệm, mô hình sử dụng hàm AddLearningResource để ghi tài nguyên và các khái niệm vào cơ sở dữ liệu đồ thị. Mỗi tài nguyên học tập được gắn với các thuộc tính như độ khó (difficulty) và độ liên quan (relevance), giúp phân loại tài nguyên học tập theo các tiêu chí quan trọng.

### **5.5.1 Cấu Trúc Đồ Thị**

Hệ thống sử dụng cơ sở dữ liệu đồ thị Neo4j để lưu trữ và quản lý các khái niệm học tập, tài nguyên học tập, và mối quan hệ giữa chúng. Cấu trúc đồ thị bao gồm các loại node và mối quan hệ sau:

- **Concept:** Mỗi node đại diện cho một khái niệm học tập.



- **Learning Resource:** Các node tài nguyên học tập đại diện cho các bài giảng, sách, video, tài liệu học tập, ví dụ: "Video bài giảng về phương trình bậc hai", "Bài viết về đạo hàm".
- **Learner:** Node người học chứa thông tin về người học, bao gồm các khái niệm mà họ đang học và mức độ thành thạo đối với các khái niệm này.

Các mối quan hệ giữa các node bao gồm:

- **COVERS:** Quan hệ giữa tài nguyên học tập và khái niệm mà tài nguyên đó đề cập đến. Mỗi quan hệ này có thuộc tính **difficulty**, đại diện cho độ khó của tài nguyên học tập đối với khái niệm đó. Ví dụ: Một video có thẻ "COVERS" khái niệm "Phương trình bậc hai" với độ khó là 0.8.
- **LEARNS:** Quan hệ giữa người học và khái niệm mà họ đang học. Mỗi quan hệ này có thuộc tính **proficiency**, biểu thị mức độ thành thạo của người học đối với khái niệm đó. Ví dụ: Người học "LEARNS" khái niệm "Lập trình Python" với mức độ thành thạo là 0.6.

### 5.5.2 Thiết kế prompt cho việc đề xuất lộ trình học

Việc đánh giá các tài nguyên học tập phù hợp sẽ được phụ trách bởi LLM. Sau đây là câu prompt đầu vào của hệ thống:

You are a teacher assistant, and your task is to recommend learning resources based on the learner's profile. Here's how the process works:

Receive the learning goal from the learner.

Invoke the function `get_learner_profile_and_learning_resources` with the list of analyzed concepts. This will return two things:

The learner's profile, which contains the proficiency level (0 to 1) for each concept. For instance:

High proficiency (0.8-1): The learner is very skilled in the concept (e.g., Linked List).

Low proficiency (0.0-0.3): The learner is not familiar with the concept (e.g., Binary Tree).

Medium proficiency (0.3-0.7): The learner has some experience but needs improvement (e.g., sorting algorithms).

A list of learning resources that match the learner's needs.



Select the top 5 most suitable resources, based on the learner's profile and the resources returned.

For each resource, explain why it is a good match for the learner:

Focus on what the learner needs to improve on and how the resource addresses those weaknesses.

Avoid mentioning numerical proficiency, but instead focus on the learner's needs.

Explain what the learner can achieve by interacting with each resource.

Call the lp\_recommender\_response tool with your analysis and recommendations.

Example:

Learning Goal: Mastering Binary Trees

Learner Profile:

Binary Trees: Proficiency = 0.3 (moderate skill)

Data Structures: Proficiency = 0.8 (high skill)

Algorithms: Proficiency = 0.5 (moderate skill)

Learning Resources:

Resource 1: "Binary Tree Fundamentals" (difficulty 0.4)

Resource 2: "Advanced Binary Tree Algorithms" (difficulty 0.7)

Resource 3: "Data Structure Optimization" (difficulty 0.6)

Resource 4: "Understanding Tree Traversals" (difficulty 0.5)

Resource 5: "Recursive Tree Operations" (difficulty 0.6)

Recommended Resources and Explanations:

Resource 1: "Binary Tree Fundamentals"

Based on your profile, you have a moderate understanding of Binary Trees, but there are still fundamental areas where you could improve. This resource will help you strengthen your foundational knowledge, ensuring that you grasp basic concepts and terminology. This is particularly important because a strong foundation in Binary Trees will make



advanced topics easier to comprehend.

**Resource 2: "Advanced Binary Tree Algorithms"**

Your current proficiency in Binary Trees shows you have a basic understanding, but there are gaps when it comes to more complex algorithms. This resource will help you expand your knowledge of how Binary Trees can be manipulated and optimized through more advanced techniques, which will enhance your overall problem-solving ability.

**Resource 3: "Data Structure Optimization"**

Since you are already proficient in data structures, this resource will help you learn how to make data structures, including Binary Trees, more efficient. Understanding optimization strategies will not only improve your technical skills but will also help you think more critically about performance in algorithms.

**Resource 4: "Understanding Tree Traversals"**

Although you have a moderate proficiency with Binary Trees, your understanding of traversal techniques may need refinement. This resource is designed to help you understand different tree traversal methods, which are crucial for manipulating and accessing data within Binary Trees. Mastering this concept will be vital for solving problems efficiently.

**Resource 5: "Recursive Tree Operations"**

Your profile suggests that you may not be fully comfortable with recursive algorithms, which are essential for operations in Binary Trees. This resource will walk you through recursive techniques and give you the confidence to handle recursive solutions within tree structures.

Quy trình xử lý mục tiêu học tập và đề xuất tài nguyên học tập được thực hiện thông qua các bước sau:

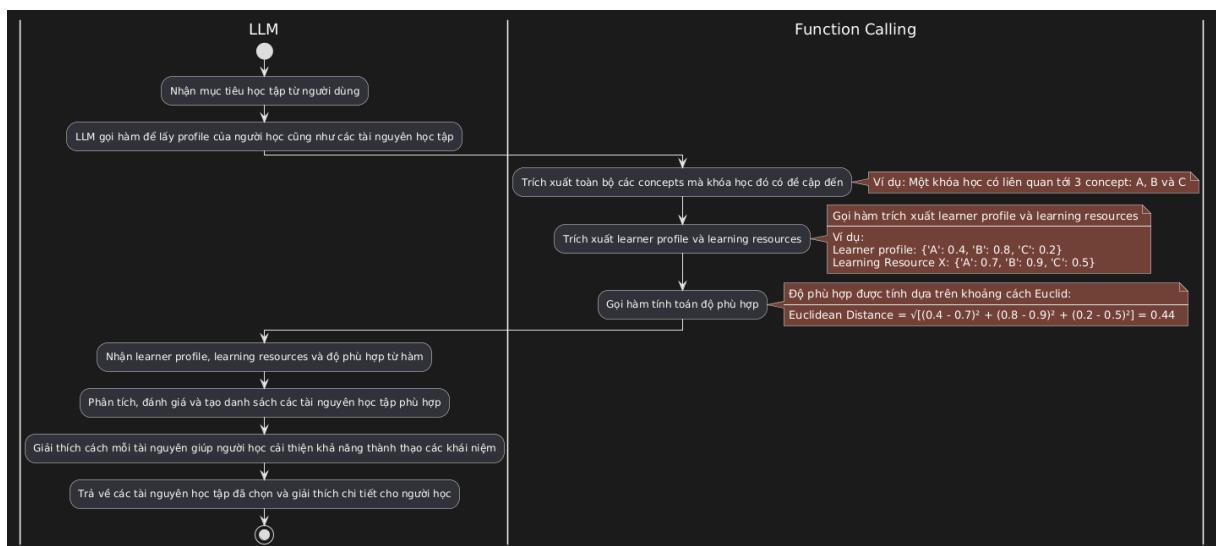
1. LLM sẽ sử dụng function calling để gọi các hàm trong hệ thống nhằm trích xuất các thông tin sau:
  - Tất cả các khái niệm liên quan mà khóa học đó có đề cập sẽ được trích xuất.

- Từ các khái niệm trên, profile của người học sẽ được mô hình thành một vector nhiều chiều, trong đó, mỗi thành phần của một vector là mức độ thông thạo của người học ứng với một khái niệm nhất định.
- Tương tự với profile của người học, các tài nguyên học tập cũng sẽ được mô hình thành các vector, trong đó, mỗi một thành phần biểu diễn độ khó của tài nguyên đó với một khái niệm nhất định.
- Từ vector biểu diễn profile của người học và vector biểu diễn các tài nguyên học tập, độ phù hợp của người học ứng với một tài nguyên học tập tương ứng sẽ được tính bằng cách Euclid. Công thức tính Euclidean distance là:

$$\text{Euclidean Distance} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Trong đó,  $x_i$  và  $y_i$  là các component của vector  $X$  và  $Y$  tương ứng, đại diện cho các khái niệm hoặc tài nguyên học tập.

2. Sau khi có được profile của người học, các tài nguyên học tập tương ứng cũng như độ phù hợp, LLM sẽ phân tích, đánh giá để chọn ra cũng như giải thích chi tiết về cách mỗi tài nguyên giúp người học cải thiện khả năng thành thạo các khái niệm cần thiết. LLM sẽ trả về các giải thích dạng tự nhiên cho người học.



**Figure 5.20: Giải thích chi tiết về các tài nguyên học tập**

## 5.6 Thiết kế Wireframe cho giao diện

### 5.6.1 Landing Page

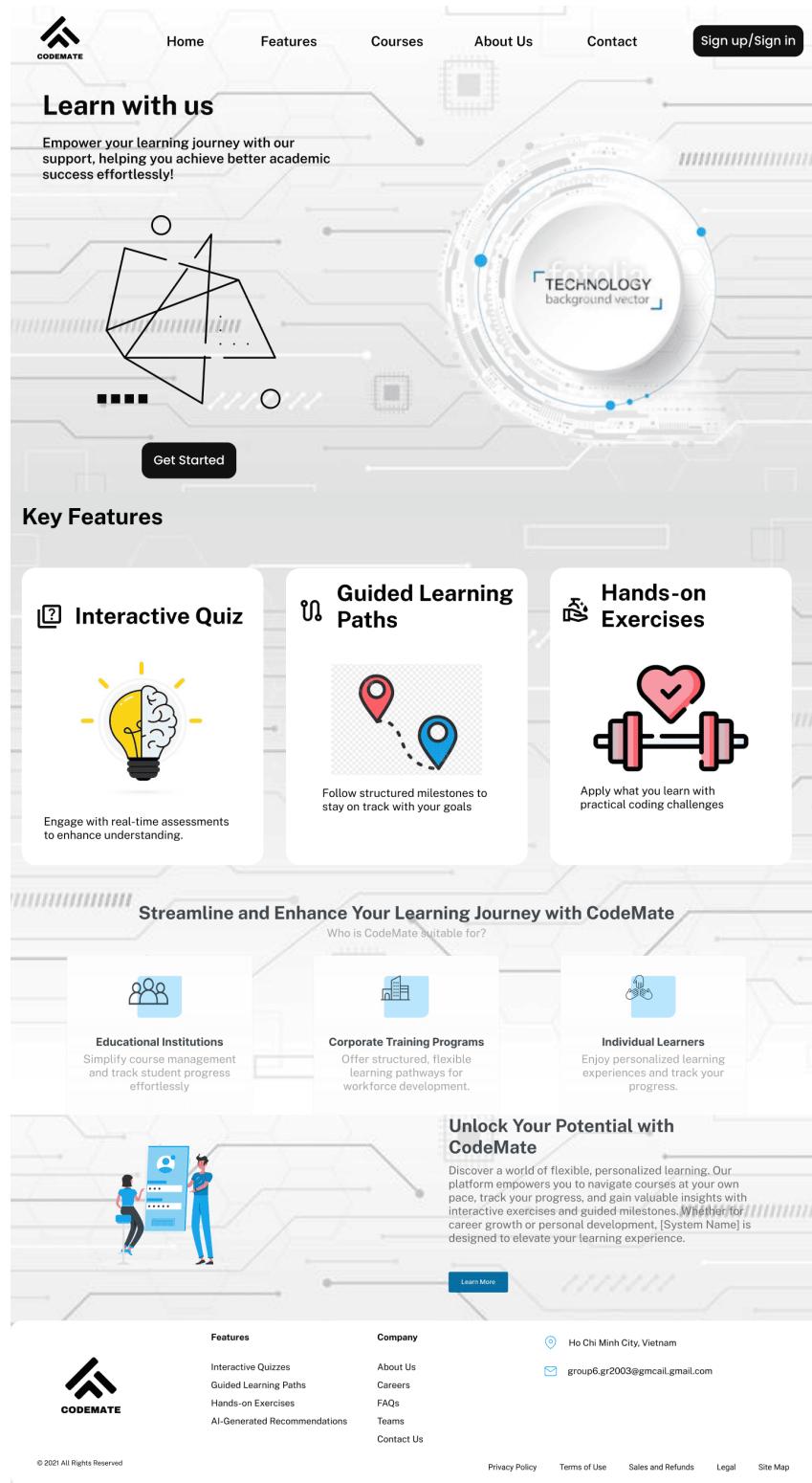


Figure 5.21: Landing Page

Landing Page mô tả tổng quát về website hỗ trợ học tập **CODEMATE**.

## 5.6.2 Authentication

### 5.6.2.1 Log In Screen

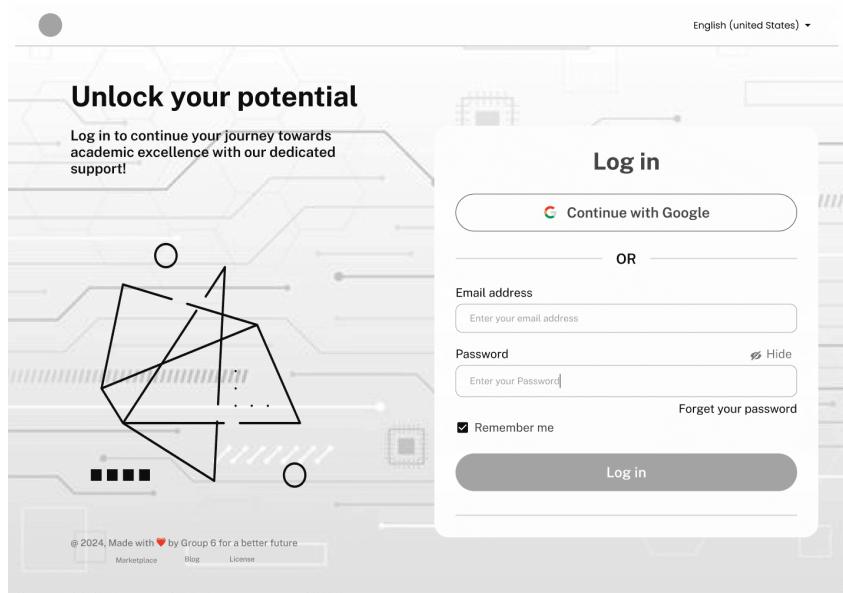


Figure 5.22: Log In Screen

Trang đăng nhập hệ thống, cho phép người dùng đăng nhập vào hệ thống bằng Email của mình hoặc qua phương thức đăng nhập bằng bên thứ ba (**Google**). Hệ thống sẽ nhận diện tên miền riêng của Email để xác định tính xác thực của Email trường Đại học do người dùng cung cấp. Nếu người dùng không dùng Email do trường Đại học cung cấp, sẽ không thể truy cập vào hệ thống. Ngoài ra, role của người dùng cũng sẽ được hệ thống phân loại dựa vào Email Address của người dùng và cơ sở dữ liệu của trường Đại học cung cấp cho hệ thống.

### 5.6.2.2 Forgot Password Screen

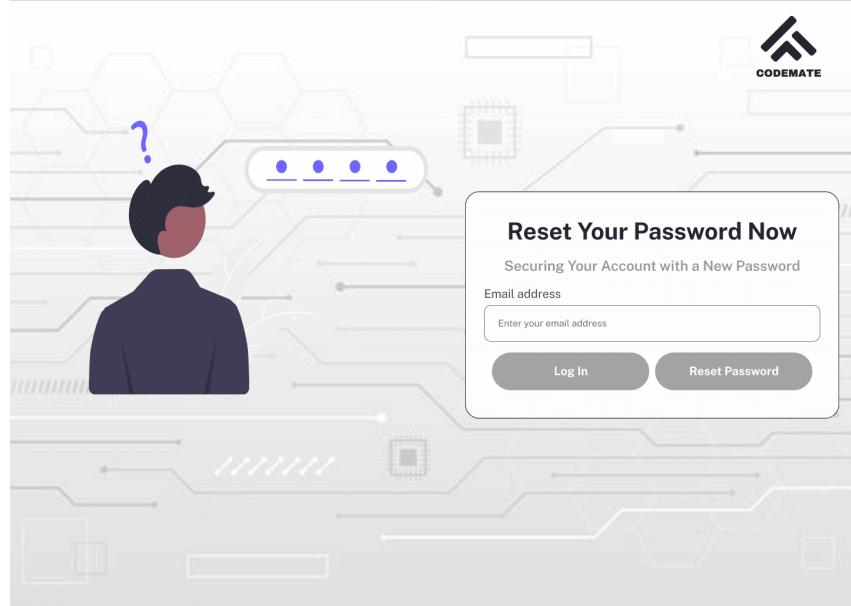


Figure 5.23: Forgot Password Screen

### 5.6.2.3 Reset Password Screen

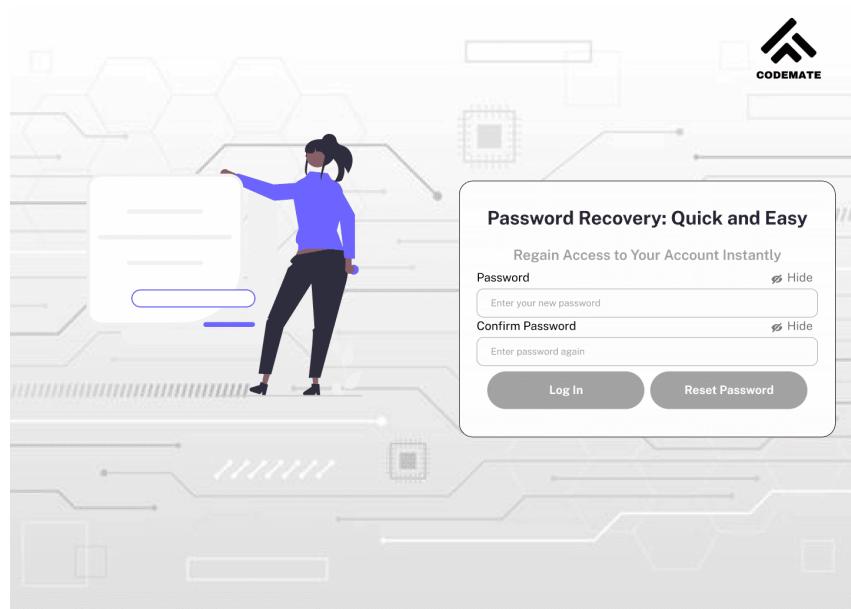


Figure 5.24: Reset Password Screen

Các trang "**Forgot Password**" và "**Reset Password**" hỗ trợ người dùng trong trường hợp quên mật khẩu truy cập vào hệ thống và muốn lấy lại mật khẩu của mình thông qua địa chỉ Email.



## 5.6.3 Course Management (Student)

### 5.6.3.1 Courses List Screen

The screenshot shows the 'CODEMATE' platform interface. On the left, there's a sidebar with a search bar, dashboard, courses (selected), notifications, progress tracking, exercises, help, and settings. The main area displays 'Recent Courses' with three cards: 'The Science of Well-Being', 'Python for Everybody Specialization', and 'The Science of Chemistry'. Below this is a section for 'My Courses' with four listed: 'Python for Everybody Specialization', 'Introduction to Data Science', 'Product Management Basic - Course', and 'BM Data Science Professional Certificate'. Each course card includes a thumbnail, title, start date, duration, student count, learning outcomes, professor name (Devi Prasad), and a 'View Course' button.

Figure 5.25: Courses List Screen

Trang danh sách khóa học, hiển thị danh sách các khóa học mà sinh viên đã đăng ký tại học kỳ đó. Đặc biệt là các khóa học liên quan đến lĩnh vực "**Khoa học máy tính**". Sinh viên có thể xem các khóa học mình đã và đang học hoặc các khóa học truy cập gần đây. Mỗi khóa học hiển thị các thông tin như: **Tên khóa học**, **Tên giảng viên**, **Thời gian bắt đầu**, **Thời gian kết thúc**, **Đầu ra môn học** và tác vụ "**Xem chi tiết khóa học**".

A modal window titled 'Learning Outcomes' is shown. It features a flag icon and a close button. The content lists five learning outcomes with icons: a red circle for 'Develop data-driven decision-making skills.', another red circle for 'Understand the basics of data science and its applications.', a green checkmark for 'Analyze and visualize data using Python and relevant libraries.', and a green checkmark for 'Apply statistical methods and machine learning algorithms to solve real-world problems.'

Figure 5.26: Modal Learning Outcomes

Sinh viên có thể xem chi tiết "**Đầu ra môn học**" bằng nút "**View more**"



### 5.6.3.2 Detailed Course Screen

Trang thông tin chi tiết khóa học, field Description, cho phép sinh viên xem mô tả "**Đầu ra môn học**" do giảng viên cung cấp. Ngoài ra, các thông tin của khóa học như: **Tên môn học**, **Phản trǎm hoàn thành hiện tại khóa học của sinh viên**, **Label trạng thái môn học**, **Số unit của khóa học**, **Tên giảng viên** được hiển thị chung ở một component. Mỗi khóa học sẽ có 3 fields hỗ trợ cho các thông tin: **Description**, **Lessons**, **Exercises**.

The screenshot shows the 'Course Details' page for a 'Data Science' course. On the left is a sidebar with navigation links: Search, Dashboard, Courses (highlighted), Notifications, Progress Tracking, Exercises, Help, and Setting. At the top right are icons for notifications, messages, and user profile (Swetha shankares). The main content area has a large blue header image with a person icon. Below it, the course title 'Data Science' is displayed with a progress bar showing '65%' completed. To the right of the progress bar are the dates '22/8/2024-25/5/2024', '10 Chapters', and the author 'Adam Levi'. A 5-star rating is shown. A 'Recommend Lessons' section features a call-to-action button with the text 'Join our Recommend Learning – tailored lessons to reach your goals, one step at a time.' and a right-pointing arrow. Below the main title, there's a 'Description' tab selected, followed by 'Lessons', 'Exercises', and 'Recommend Lessons'. The 'Description' tab contains a list of learning outcomes:

- Understand the basics of data science and its applications.
- Analyze and visualize data using Python and relevant libraries.
- Apply statistical methods and machine learning algorithms to solve real-world problems.
- Develop data-driven decision-making skills.

Figure 5.27: Detailed Course Screen - Description

Field Description chứa thông tin liên quan đến mô tả khóa học và đầu ra môn học do giảng viên cung cấp.

This screenshot shows the 'Lessons' section of the 'Data Science' course page. The sidebar and top navigation are identical to Figure 5.27. The main content area displays three chapters under the 'Lessons' tab:

Chapter 1: Introduction to Data Science	5 documents	Download	Print	Share
Covers techniques for handling missing data, outliers, and transforming data for analysis.	5 documents	Download	Print	Share
Covers techniques for handling missing data, outliers, and transforming data for analysis.	5 documents	Download	Print	Share
Covers techniques for handling missing data, outliers, and transforming data for analysis.	5 documents	Download	Print	Share

Figure 5.28: Detailed Course Page - Lessons



Field Lessons chứa các lessons được chia bởi giảng viên. Mỗi lesson được hiển thị với thông tin: **Tên lesson và Số tài liệu tham khảo được giảng viên đăng tải**

The screenshot shows the 'Data Science' course details page. On the left sidebar, there are links for Dashboard, Courses (selected), Notifications, Progress Tracking, Exercises (selected), Help, and Setting. The main content area displays the course title 'Data Science', completion percentage '65%', and a progress bar indicating 'Completed'. It also shows the start date '22/8/2024 - 25/5/2024', number of chapters '10 Chapters', and author 'Adam Levi'. Below this, there are five exercises listed:

Description	Due Date	Points	Action
Assignment 1 Data Science	30/10/2024 23:59:59	100 points	Start
Quiz 1: Create Plan Data Science	N/A	100 points	Start
Quiz 2 Data Science	30/10/2024 23:59:59	100 points	Start
Quiz 3 Data Science	N/A	100 points	Start
Quiz 4 Data Science	30/10/2024 23:59:59	100 points	Start

Figure 5.29: Detailed Course Page - Exercises

Field Exercises chứa các bài Quiz/Assignment do giảng viên hoặc hệ thống cung cấp hỗ trợ cho lesson/khoa học đó.

The screenshot shows the 'Data Science' course details page. The sidebar and course summary are identical to Figure 5.29. The main content area displays four recommended lessons:

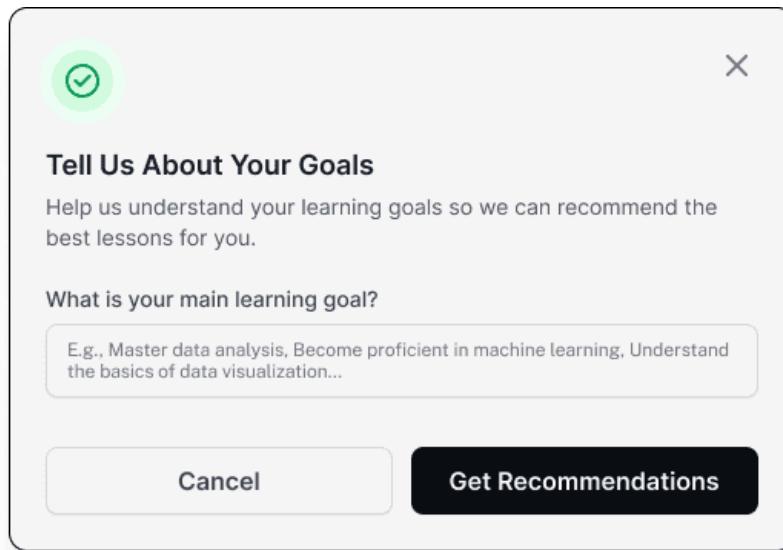
Lesson Description	Actions
Chapter 3: Introduction to Data Science Covers techniques for handling missing data, outliers, and transforming data for analysis.	View, Edit, Delete, Preview
Chapter 5: Introduction to Data Science Covers techniques for handling missing data, outliers, and transforming data for analysis.	View, Edit, Delete, Preview
Chapter 7: Introduction to Data Science Covers techniques for handling missing data, outliers, and transforming data for analysis.	View, Edit, Delete, Preview
Chapter 8: Introduction to Data Science Covers techniques for handling missing data, outliers, and transforming data for analysis.	View, Edit, Delete, Preview

Figure 5.30: Detailed Course Page - Recommend Lessons

Field Recommend Lessons Chứa các lessons mà hệ thống đã đề xuất cho sinh viên.



#### 5.6.3.3 Define Goal Modal



**Figure 5.31:** Modal để lấy mục tiêu của sinh viên

Sinh viên có thể cung cấp mục tiêu học tập của mình thông qua modal ở trang chi tiết khóa học

#### 5.6.3.4 Modules

Sau khi chọn mục tiêu học tập và được hệ thống đề xuất lesson, sinh viên sẽ được truy cập vào lesson được đề xuất với giao diện chi tiết lesson như sau:



The screenshot shows a user profile for 'Swetha shankaresh' (Student) on the left. The main content area displays a lesson titled 'Chapter 1: Introduction to Data Science'. Key details include a progress bar at 65%, a recommended time of 1.5 hours, and 4 modules completed. A 'Completed' badge is present. To the right, 'Learning Outcomes' are listed:

- After completing this lesson, you'll be able to:
- Explain the process of semantic analysis.
- Perform type checking and type inference.
- Implement symbol tables and manage variable scopes.
- Handle errors that arise during semantic analysis.

Below this, 'Recommend Content:' is described as focusing on semantic analysis, lexical and syntax analysis, and symbol tables. 'Explain:' notes the user's strong foundation in lexical and syntax analysis, which adds meaning to the parsed program. 'Modules:' are listed as 'Check for type compatibility', 'Verify variable declarations', and 'Ensure followed scope rules'.

Figure 5.32: Chi tiết Lesson

Lesson đề xuất sẽ cung cấp đầy đủ thông tin về **Learning Objectives**, **Recommend Content**, **Lời giải thích cho sự đề xuất và các modules cần học**

#### 5.6.3.5 Choices of Learning Styles

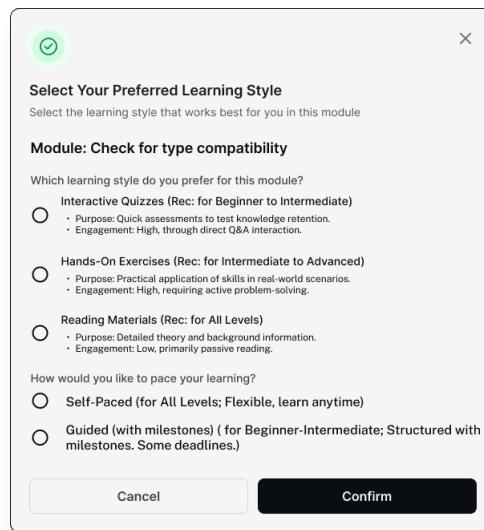


Figure 5.33: Modal cho phép sinh viên lựa chọn kiểu học phù hợp



Sinh viên cung cấp thông tin về cách học mong muốn và tốc độ phù hợp với bản thân ở Modal này để quyết định cách học cho từng module

The screenshot shows the CODEMATE learning platform interface. On the left is a sidebar with navigation links: Dashboard, Courses (which is selected and highlighted in grey), Notifications, Progress Tracking, Exercises, Help, and Setting. At the bottom of the sidebar is a user profile for 'Swetha shankaresh' with the title 'Student'. The main content area has a search bar at the top right. Below it, a 'Lesson Details' section shows the title 'Self-Paced with Interactive Quiz Learning Style: "Check for Type Compatibility"'. A welcome message reads: "Welcome to the 'Check for Type Compatibility' module! This module uses interactive quizzes to help you test your knowledge. Work through each quiz at your own speed, and check your answers along the way. You can retake any quiz for better understanding." To the right of the message is a 'Learning Objectives' section with three bullet points: 'Recognize type compatibility rules.', 'Apply type conversion methods.', and 'Solve common type compatibility problems.' A large 'Do Quiz' button is located below the message. At the very bottom of the main content area is another user profile for 'Swetha shankaresh'.

**Figure 5.34:** Giao diện khi sinh viên chọn học theo cách học làm Quiz với tiến độ tự do



The screenshot shows the CodeMate platform interface. On the left is a sidebar with icons for Dashboard, Courses (which is selected), Notifications, Progress Tracking, Exercises, Help, and Setting. At the bottom of the sidebar is a user profile for 'Swetha shankaresh Student'. The main content area has a search bar and a 'Lesson Details' header. Below it is a section titled 'Self-Paced with Reading Materials: "Check for Type Compatibility"'. It includes a welcome message, learning objectives, and reflection questions. A code example is shown:

```
interface Person {  
    name: string;  
}  
  
let person: Person = { name: "Alice" };  
let anotherPerson = { name: "Bob", age: 30 }; // Additional properties are allowed  
  
person = anotherPerson; // Compatible because 'anotherPerson' has at least 'name'.
```

The text below the code states: 'Here, anotherPerson is compatible with Person even though it has an extra property (age), because it matches the required structure of Person.'

Figure 5.35: Giao diện khi sinh viên chọn học theo cách học đọc tài liệu với tiến độ tự do

The screenshot shows the CodeMate platform interface. The sidebar and user profile are identical to Figure 5.35. The main content area has a search bar and a 'Lesson Details' header. Below it is a section titled 'Self-Paced with Hands-On Exercises: "Check for Type Compatibility"'. It includes a welcome message, learning objectives, and an 'Exercises List' section with two tasks:

1. Write a function that takes an any type parameter and returns it as a number. Use type assertion.
2. The following code has a type compatibility error. Identify and fix it.

A 'Do all exercises' button is located at the bottom of the exercises list.

Figure 5.36: Giao diện khi sinh viên chọn học theo cách học làm bài tập Coding với tiến độ tự do

## 5.6.4 User Profile

### 5.6.4.1 Profile Information

The screenshot shows the 'My profile' page with a sidebar containing links for Dashboard, Courses, Notifications, Progress Tracking, Help, and Setting. The main area is titled 'USER INFORMATION' with fields for Username, Email address, First name, Last name, Birth Year, Organization, and Phone Number. It includes a 'Change Image' button next to a placeholder profile picture and a 'Change Password' button.

(a) Trang thông tin tài khoản

The screenshot shows the same 'My profile' page as above, but the 'Edit' button at the top right is highlighted in blue, indicating the user is in edit mode to update their profile information.

(b) Chính sửa thông tin tài khoản

Trang "Thông tin tài khoản" cung cấp các trường thông tin cá nhân của người dùng bao gồm: Tên người dùng (Username), Địa chỉ email (Email address), Tên (First name), Họ (Last name), Năm sinh (Birth Year), Trường (Organization), và Số điện thoại (Phone Number). Người dùng có thể thay đổi hình ảnh đại diện bằng cách nhấp vào nút Change Image và cập nhật mật khẩu thông qua nút Change Password. Tất cả thông tin có thể được chỉnh sửa bằng cách nhấn vào nút Edit.

### 5.6.4.2 Change Password, Avatar

The screenshot shows the 'My profile' page with a sidebar containing links for Dashboard, Courses, Notifications, Progress Tracking, Help, and Setting. The main area is titled 'USER INFORMATION' with fields for Old Password, New Password, and Confirm Password. It includes a 'Change Image' button next to a placeholder profile picture and a 'Change Password' button.

(a) Thay đổi mật khẩu

The screenshot shows the 'My profile' page with a sidebar containing links for Dashboard, Courses, Notifications, Progress Tracking, Help, and Setting. The main area is titled 'USER INFORMATION' with a placeholder for a profile picture and a 'Choose File' button. A message says 'Please upload square image size less than 30MB'. It includes a 'Save' button and a 'Change Image' button.

(b) Thay đổi avatar

Trang "Đổi mật khẩu" cho phép người dùng thay đổi mật khẩu bằng cách nhập mật khẩu hiện tại, mật khẩu mới, và xác nhận lại mật khẩu mới. Sau khi điền đầy đủ thông tin, người dùng nhấn nút "Đổi mật khẩu" để hoàn tất. Nếu có lỗi hoặc thành công, hệ thống sẽ hiển thị thông báo tương ứng. Trang "Đổi ảnh đại diện" cho phép người dùng tải lên ảnh mới từ thiết bị của mình và nhấn "Lưu thay đổi" để cập nhật ảnh đại diện.



## 5.6.5 Dashboard Screen

The dashboard is divided into several sections:

- Left Sidebar:** Labeled "CODEMATE". Includes a search bar ("Search...") and links to "Dashboard", "Courses", "Notifications", "Progress Tracking", "Exercises", "Help", and "Setting".
- Welcome Area:** Displays a welcome message for "Swetha shankaresh", stating "Glad to see you again! Keep up the momentum in your Data Science Basics course." It also includes a "Need help?" button and a "Send Feedback" button.
- Weekly Progress Summary:** Shows a circular progress bar at 65% completion for the "Data Science" category, with a note that 2 lessons are completed.
- Courses Section:** A table titled "Courses" showing three entries:

Name	Members	COMPLETION
Chakra Soft UI Version	3 members	60%
Add Progress Track	2 members	10%
Fix Platform Errors	2 members	100%

Figure 5.39: Dashboard - Student

Trang Dashboard của sinh viên, mô tả tóm tắt về các khóa học mà sinh viên đã và đang tham gia và một số thông tin về khóa học vừa học ở lần truy cập gần nhất.



## 5.6.6 Exercise Practice

### 5.6.6.1 Code Exercise

```
// Implements
class Solution {
public:
    vector<int> twoSum(vector<int>& nums, int target) {
    }
}

Given an array of integers nums and an integer target, return indices of the two numbers such that they add up to target. You may assume that each input would have exactly one solution, and you may not use the same element twice. You can return the answer in any order.
```

**Example 1:**  
Input: `nums` = [2,7,11,15], `target` = 9  
Output: [0,1]  
Explanation: Because `nums[0]` + `nums[1]` == 9, we return [0, 1].

**Example 2:**  
Input: `nums` = [3,2,4], `target` = 6  
Output: [1,2]

**Constraints:**  
•  $2 \leq \text{nums.length} \leq 10^4$

**AI Tutor**

**Testcase**

Case1 Case2 Case3

nums = [2, 7, 11, 5]  
target = 6

Figure 5.40: Giao diện thực hành code

**Description** **Submission**

Status	Language	Runtime	Memory
Accepted	C++	3ms	14.3MB
Oct. 27/2024			
Compile Error	C++	N/A	N/A
Oct. 27/2024			
Compile Error	C++	N/A	N/A
Oct. 27/2024			

**Code**

```
// Implements
class Solution {
public:
    vector<int> twoSum(vector<int>& nums, int target) {
    }
}
```

**AI Tutor**

**Testcase**

Case1 Case2 Case3

nums = [2, 7, 11, 5]  
target = 6

Figure 5.41: Lịch sử submission

## 5.6.7 AI Tutor

Giao diện AI Tutor được thiết kế gồm ba chức năng chính như sau: cung cấp gợi ý (Get Hints), sửa lỗi (Fix Bug), và giải thích code (Explain Code).

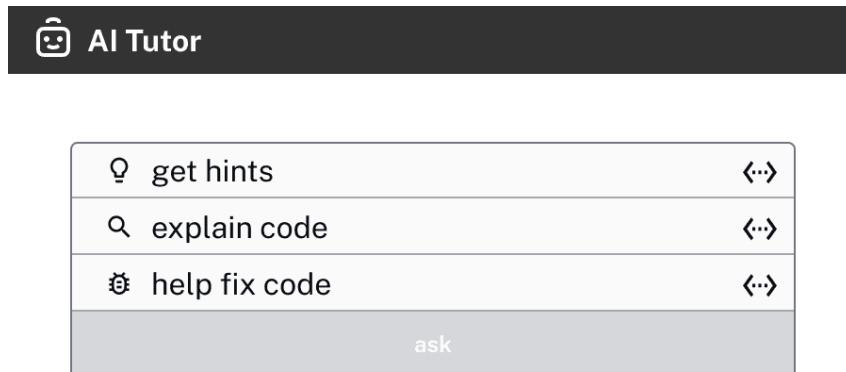


Figure 5.42: Các tính năng của AI Tutor

Trong phần giải thích code, người dùng có thể chọn đoạn mã muốn hiểu rõ hơn; AI Tutor sẽ cung cấp giải thích tổng quát cho đoạn mã đó, và khi người dùng di chuột lên từng dòng, một popup sẽ hiển thị giải thích chi tiết cho từng dòng.

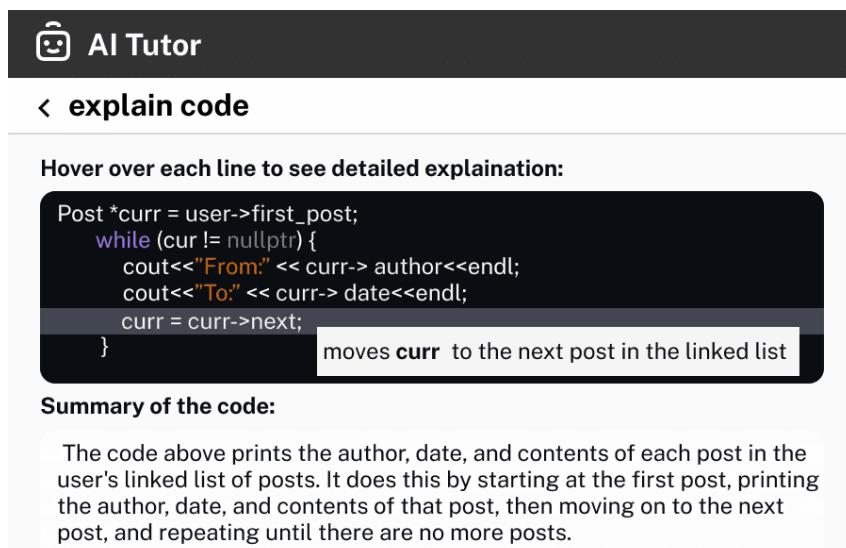


Figure 5.43: Giải thích từng dòng code

Đối với chức năng sửa lỗi, khi người dùng gặp lỗi trong quá trình compile code, AI Tutor sẽ phân tích và giải thích chi tiết nguyên nhân lỗi, đồng thời đề xuất cách khắc phục, giúp người dùng học hỏi từ lỗi sai của mình.



AI Tutor

< help fix code

```
vector<int> numbers = {1, 2, 3, 4, 5};  
for (int i = 0; i <= numbers.size(); i++) {  
    cout << numbers[i] << " ";  
}
```

**Fixes:** The error occurs on the line `for (int i = 0; i <= numbers.size(); i++)`. When `i` equals `numbers.size()`, the program will try to access an element outside the bounds of the vector `numbers`, causing a runtime error.

**Suggested Fixes:** Change the loop condition: Modify the condition `<=` to `<` to ensure the loop only accesses elements within the valid range of `numbers`.

Figure 5.44: Hỗ trợ sửa lỗi

Khi người dùng cần gợi ý cho bài tập, họ có thể nhấn nút "Get Hints" để nhận những hướng dẫn hữu ích; nếu chưa rõ, AI Tutor còn có thể đưa ra ví dụ minh họa giúp họ hiểu rõ hơn.

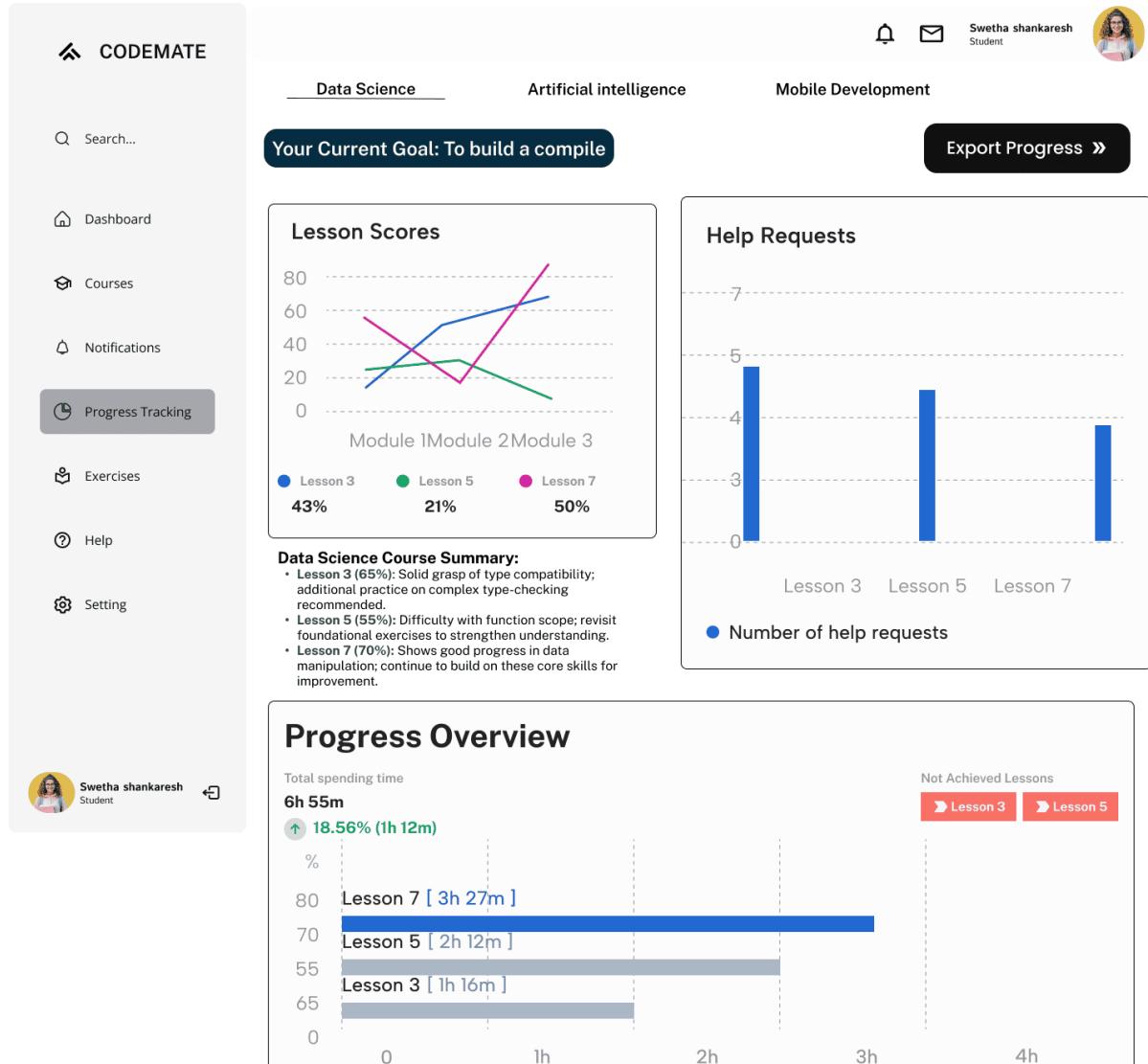
AI Tutor

< get hints

1. use a loop to iterate over the string
2. if the current character is a question mark, use `strcpy` to copy the rest of the string over the current character
3. use `strlen` to get the length of the string
4. use `realloc` to resize the string

Figure 5.45: Cung cấp gợi ý cho bài tập

## 5.6.8 Progress Tracking



**Figure 5.46: Progress Tracking Page**

Trang Progress Tracking ghi lại tiến độ học tập của sinh viên theo từng môn học, bao gồm 3 biểu đồ chính:

### 1. Biểu Đồ Đánh Giá Điểm Số Theo Lesson (Biểu đồ đường)

- **X-Axis:** Tên module (1, 2, 3).
- **Y-Axis:** Điểm số (từ 0 đến 100%).
- **Màu của đường biểu đồ:** Lesson(3,5,7)
- **Đơn vị:** Phân trăm (%).
- **Ý Nghĩa:** Biểu đồ này cho phép sinh viên thấy tỷ lệ hoàn thành mục tiêu của từng



module theo lesson và dễ dàng nhận ra lesson nào đạt hoặc chưa đạt yêu cầu, bị khó khăn ở module nào.

## **2. Biểu Đồ Theo Dõi Số Lượng Vấn Đề hoặc Yêu Cầu Trợ Giúp (Biểu đồ cột)**

- **X-Axis:** Tên lesson (1, 2, 3).
- **Y-Axis:** Số lần sinh viên yêu cầu trợ giúp (Đếm số lần yêu cầu trợ giúp tới AI).
- **Đơn vị:** Lần (Số lần hỏi).
- **Ý Nghĩa:** Biểu đồ này ghi nhận độ khó của từng bài tập, giúp hệ thống nhận biết các phần cần được cải thiện hoặc các bài tập khó để review lại cho giáo viên hoặc sinh viên cần chú trọng học lại.

## **3. Biểu Đồ Tổng Quan Tiến Độ Học (Progress Overview) (Biểu đồ cột)**

- **X-Axis:** Thời gian (Số giờ học).
- **Y-Axis:** Tỷ lệ hoàn thành (%) của lesson.
- **Đơn vị:** Phần trăm (%) hoàn thành lesson.
- **Ý Nghĩa:** Cho thấy tiến độ tổng thể của sinh viên, giúp họ nhận ra tốc độ học của mình và đối chiếu với lộ trình học ban đầu.

# Chương 6

## Hiện thực và Kiểm thử

### 6.1 Hiện thực hệ thống

Hệ thống học tập được triển khai nhằm hỗ trợ sinh viên và giảng viên trong việc tổ chức và quản lý việc học tập hiệu quả hơn. Với sự hỗ trợ của Mô hình Ngôn ngữ Lớn (LLMs), hệ thống có thể tự động tạo ra các bài học đề xuất dựa trên mục tiêu học tập của sinh viên và nội dung được cung cấp bởi giảng viên.

Hệ thống được thiết kế với kiến trúc phân tách rõ ràng giữa frontend và backend, sử dụng các công nghệ hiện đại để đảm bảo tính linh hoạt, dễ mở rộng và bảo trì.

#### Frontend

Frontend được xây dựng bằng VueJS, tích hợp Vuetify và Tailwind CSS để cung cấp giao diện người dùng trực quan, hiện đại và dễ sử dụng. Các chức năng chính bao gồm:

- Quản lý dashboard của sinh viên
- Hiển thị danh sách khóa học, chi tiết khóa học
- Gửi phản hồi và nhập mục tiêu học tập

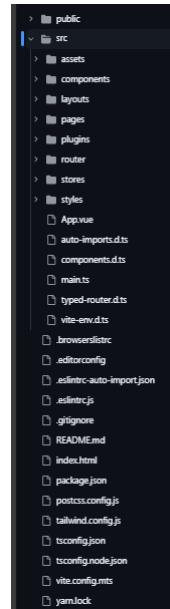


Figure 6.1: Cấu trúc Frontend

## Hệ thống thiết kế (Design System)

Hệ thống thiết kế được triển khai với mục tiêu cung cấp một giao diện người dùng nhất quán, dễ sử dụng và dễ mở rộng. Hệ thống này bao gồm các thành phần chính sau:

### Typography

Typography được xây dựng trên cơ sở font chữ Public Sans, với nhiều kích cỡ và độ đậm khác nhau để phục vụ các nhu cầu hiển thị khác nhau. Font chữ này được chọn vì dễ đọc, phù hợp với các thiết kế hiện đại và hỗ trợ đa ngôn ngữ. Các cấp độ bao gồm:

- **Heading:** Từ Heading 1 (48px, Bold) đến Heading 4 (24px, Semibold).
- **Body Text:** Các cấp độ từ Large đến Small, sử dụng độ đậm từ Regular đến Bold, kích thước từ 10px đến 20px.



Typography		
<b>Public Sans</b>		
Font Public Sans dễ đọc, dễ nhìn, phù hợp với thiết kế hiện đại và sạch sẽ, phù hợp với nhiều loại dự án, ngoài ra font cũng hỗ trợ nhiều ngôn ngữ		
<b>Heading 1</b>	<b>Bold</b>	<b>48px</b>
<b>Heading 2</b>	<b>Bold</b>	<b>40px</b>
<b>Heading 3</b>	<b>Bold</b>	<b>32px</b>
Heading 4	Semibold	24px
Body/Large 1	<b>Bold</b>	20px
Body/Large 2	Semibold	20px
Body/Large 3	Medium	20px
Body/Large 4	Regular	20px
Body/Base 1	<b>Bold</b>	16px
Body/Base 2	Semibold	16px
Body/Base 3	Medium	16px
Body/Base 4	Regular	16px
Body/Small 1	<b>Bold</b>	12px
Body/Small 2	Semibold	12px
Body/Small 3	Medium	12px
Body/Small 4	Regular	12px
Body/Ex-Small 1	<b>Bold</b>	10px
Body/Ex-Small 2	Semibold	10px
Body/Ex-Small 3	Medium	10px
Body/Ex-Small 4	Regular	10px

Figure 6.2: Hệ thống Typography sử dụng trong ứng dụng

### Bảng màu (Color Palette)

Bảng màu được thiết kế để mang lại sự nhất quán và dễ nhận biết trong giao diện người dùng. Bảng màu chính bao gồm các màu **Primary**, **Secondary**, và các trạng thái như **Error**, **Warning**, **Neutral**, **Success**. Mỗi màu được chia thành các sắc độ từ nhạt (25, 50) đến đậm (900, 950), nhằm đảm bảo khả năng hiển thị tốt trong các trường hợp sử dụng khác nhau.

- **Primary:** Dùng cho các thành phần chính như nút bấm và tiêu đề.
- **Secondary:** Dùng cho các chi tiết phụ hoặc nhấn mạnh nội dung.
- **Error, Warning, Success:** Dùng để phản hồi trạng thái của hệ thống.



Primary	25 #e0d1ff	50 #a5c9ff	100 #6587ff	200 #8677ff	300 #7950ff	400 #6c3cd2	500 #2000ee	600 #5400ff				
Primary Variant	600 #3700b3	700 #4700f1	800 #3a0093	900 #200077	950 #21005c							
Secondary	25 #7fe6d7	50 #87e2d1	100 #48dec2	200 #03dac6	400 #022bdac	600 #02a292	800 #01877a	900 #016c62	950 #00544b			
Secondary Variant	600 #018786	800 #01777a	900 #016c62	950 #00544b								
Error	25 #e0a8a3	50 #e0928d	100 #d97777	200 #e06662	300 #e0644d	400 #e05237	500 #e05020	600 #e0501a	700 #e02015	800 #e0c00f	900 #e0800a	
Warning	AA 5.28 #FFFFCF	AA 5.17 #FFFACB	AA 4.75 #FFEFCC	4.15 #FECB89	1.54 #FECB48	1.84 #FEB222	2.34 #F79009	3.49 #DC8803	AA 5.4 #D8470B	AAA #93370D	AAA #7A2E0E	AAA #4E1D09
Neutral	25 #e0e0e0	50 #d0d0d0	100 #c0c0c0	200 #b0b0b0	300 #a0a0a0	400 #909090	500 #808080	600 #707070	700 #606060	800 #505050		
Success	AA 5.31 #F0F0F0	AA 5.17 #E0E0E0	AA 4.79 #D0D0D0	4.24 #B0B0B0	1.52 #A0A0A0	1.91 #909090	2.63 #808080	3.74 #707070	AA 5.44 #606060	AAA #505050	AAA #404040	AAA #303030

Figure 6.3: Bảng màu được sử dụng trong ứng dụng

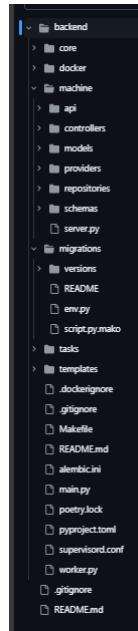
Hệ thống này được tích hợp với Tailwind CSS, giúp tự động hóa việc áp dụng các quy tắc thiết kế (design tokens) vào mã nguồn. Điều này làm giảm thiểu lỗi thiết kế, đồng thời đảm bảo giao diện nhất quán trong toàn bộ ứng dụng.

## Backend

Backend được phát triển với FastAPI, kết hợp với PostgreSQL để lưu trữ dữ liệu. Kiến trúc hệ thống được tổ chức theo các thành phần chính:

- Repository
- Provider
- Controller

Mỗi thành phần đảm nhiệm một vai trò cụ thể, giúp hệ thống trở nên dễ hiểu, có tính tổ chức cao và dễ bảo trì.



**Figure 6.4: Cấu trúc Backend**

## 6.2 Kiến trúc Backend

### 1. Repository

Repository chịu trách nhiệm quản lý các thao tác liên quan đến dữ liệu trong hệ thống. Đây là nơi thực hiện các logic liên quan đến việc truy vấn, lưu trữ và cập nhật dữ liệu từ cơ sở dữ liệu PostgreSQL.

Đặc điểm chính:

- Tách biệt logic truy xuất dữ liệu khỏi các thành phần khác
- Hỗ trợ các thao tác CRUD (Create, Read, Update, Delete)

Ví dụ: Repository sẽ xử lý các tác vụ như:

- Truy vấn thông tin sinh viên và giảng viên
- Cập nhật tiến độ học tập
- Xóa thông tin khóa học khi cần thiết

### 2. Provider

Provider chịu trách nhiệm xử lý các logic nghiệp vụ (business logic). Đây là nơi các phép toán và xử lý phức tạp được thực hiện trước khi trả kết quả cho Controller hoặc người dùng.



Đặc điểm chính:

- Không phụ thuộc vào giao diện hoặc nguồn dữ liệu cụ thể
- Tập trung vào logic nghiệp vụ và các phép toán quan trọng

Ví dụ: Provider sẽ:

- Phân tích mục tiêu học tập của sinh viên để đề xuất bài học
- Tự động tạo bài học dựa trên nội dung được giảng viên cung cấp

### **3. Controller**

Controller là nơi tiếp nhận các yêu cầu từ người dùng thông qua API, sau đó phối hợp với Provider và Repository để thực hiện các chức năng tương ứng.

Đặc điểm chính:

- Giao tiếp trực tiếp với frontend thông qua các endpoint API
- Chuyển đổi dữ liệu giữa frontend và các thành phần backend

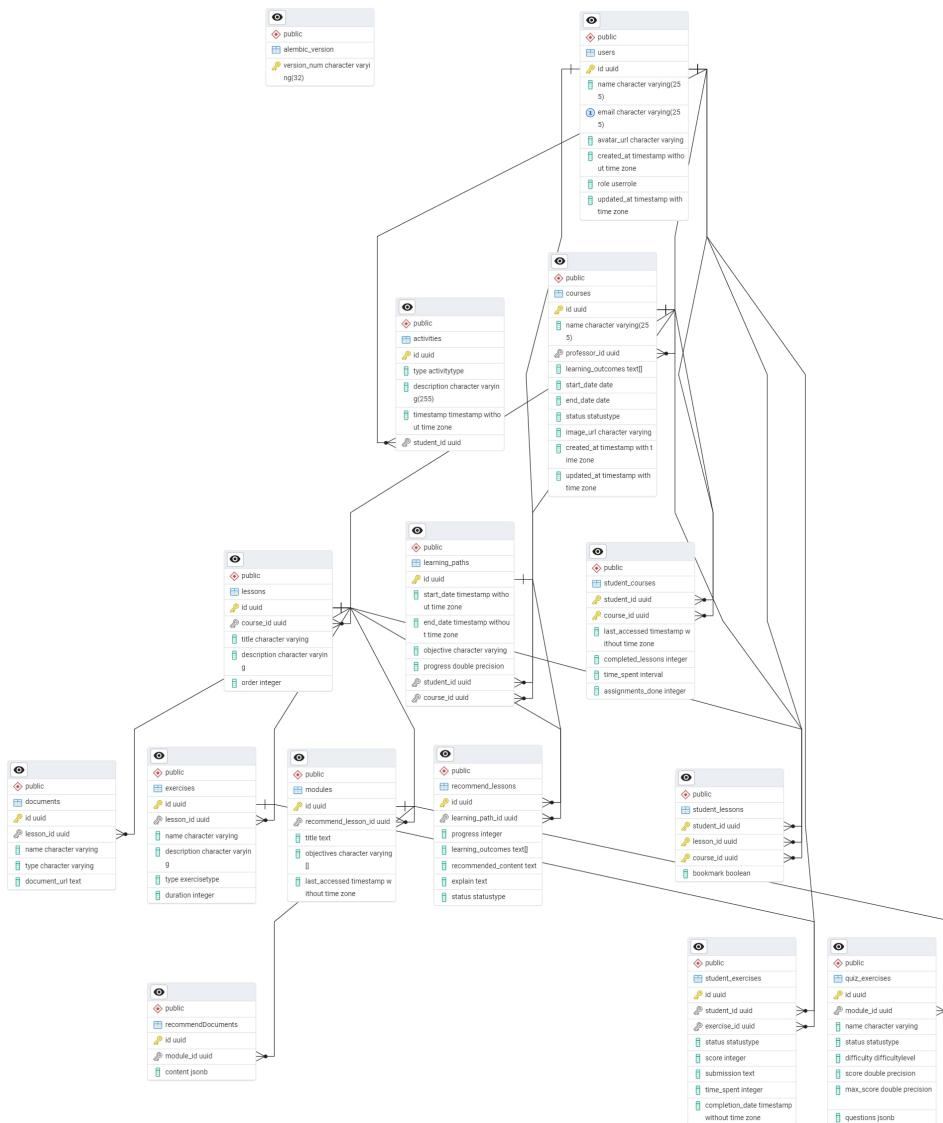
Ví dụ: Controller xử lý:

- Yêu cầu đăng nhập và xác thực người dùng
- Truy xuất danh sách khóa học và chi tiết khóa học
- Nhận mục tiêu học tập từ sinh viên và gửi đến Provider để xử lý

## **6.3 Quá trình triển khai hệ thống**

### **1. Tạo cấu trúc cơ sở dữ liệu**

Cơ sở dữ liệu được thiết kế để lưu trữ các thông tin liên quan đến sinh viên, giảng viên, khóa học và bài học. Sơ đồ dưới đây minh họa cấu trúc cơ sở dữ liệu:



**Figure 6.5: Cấu trúc Cơ sở dữ liệu**

## 2. Xây dựng API

Các API hỗ trợ giao tiếp giữa frontend và backend, đảm bảo rằng người dùng có thể tương tác với hệ thống một cách hiệu quả. Các chức năng chính của API bao gồm:

- Quản lý khóa học
- Quản lý bài học
- Gửi và nhận phản hồi từ sinh viên

FastAPI cung cấp cơ chế xử lý yêu cầu nhanh chóng và an toàn, giúp cải thiện hiệu năng hệ thống.



### **3. Quản lý bài học và đề xuất học tập**

Hệ thống sử dụng LLMs để phân tích dữ liệu từ sinh viên và giảng viên, từ đó đưa ra các bài học phù hợp. Điều này giúp cá nhân hóa việc học tập và tăng hiệu quả của hệ thống.

Cụ thể:

- Giảng viên cung cấp nội dung khóa học và mục tiêu học
- Sinh viên nhập mục tiêu học tập cá nhân
- LLMs xử lý và đề xuất các bài học tương thích

### **4. Triển khai hệ thống**

**Frontend:** Frontend của hệ thống được triển khai trên nền tảng **Vercel**, mang đến một môi trường ổn định, tối ưu cho việc triển khai các ứng dụng web. Vercel cung cấp khả năng tự động build và deploy mỗi khi có thay đổi trong mã nguồn, giúp quá trình cập nhật trở nên nhanh chóng và mượt mà. Nhờ vào quy trình này, việc triển khai và duy trì hệ thống frontend trở nên đơn giản và hiệu quả.

Vercel không chỉ đảm bảo hiệu suất cao mà còn hỗ trợ các tính năng tối ưu hóa tự động, giúp giảm thời gian tải trang và nâng cao trải nghiệm người dùng. Với khả năng mở rộng linh hoạt, hệ thống có thể dễ dàng đáp ứng với nhu cầu tăng trưởng trong tương lai.

**Link deploy frontend:** <https://codemate-fe-murex.vercel.app/>

**Backend:** Hệ thống sử dụng nền tảng **Render** để triển khai backend. Render cung cấp một môi trường chạy ổn định, dễ mở rộng và bảo mật cao, giúp tối ưu hóa việc quản lý và triển khai ứng dụng. Với quy trình build và deploy tự động, việc cập nhật hệ thống trở nên mượt mà và hiệu quả, giảm thiểu các vấn đề phát sinh trong quá trình triển khai.

Để triển khai hệ thống, chúng tôi sử dụng công cụ **Poetry** để quản lý các gói phụ thuộc (dependencies) của Python. Điều này giúp đảm bảo rằng các thư viện và công cụ cần thiết luôn được cài đặt đúng phiên bản, đồng thời giảm thiểu các lỗi liên quan đến môi trường chạy.

Hệ thống được cấu hình để tự động khởi chạy khi có yêu cầu, đảm bảo hiệu suất cao và khả năng mở rộng dễ dàng khi cần thiết. Ngoài ra, Render cũng hỗ trợ giám sát và quản lý môi trường ứng dụng, giúp dễ dàng phát hiện và xử lý các sự cố.

**Link deploy backend:** <https://edumind.onrender.com>



## 6.4 Kiểm thử

### 6.4.1 Tổng quan về kiểm thử

Mục đích của kiểm thử trong hệ thống: Đảm bảo rằng các tính năng hoạt động đúng, hệ thống không có lỗi và có thể chịu được tải trong quá trình sử dụng thực tế.

#### 6.4.1.1 Loại kiểm thử thực hiện

- **Kiểm thử chức năng:** Đảm bảo các chức năng của hệ thống hoạt động đúng
- **Kiểm thử hiệu suất:** Đảm bảo hệ thống có thể xử lý nhiều yêu cầu đồng thời
- **Kiểm thử bảo mật:** Đảm bảo hệ thống bảo vệ tốt các dữ liệu người dùng
- **Kiểm thử tích hợp:** Kiểm tra sự hoạt động đồng bộ của các module
- **Kiểm thử hồi quy:** Đảm bảo những thay đổi mới không làm hỏng các tính năng cũ

### 6.4.2 Kiểm thử API

#### 6.4.3 Mô tả chung về các kịch bản kiểm thử

Mỗi kịch bản kiểm thử được tổ chức và theo dõi theo bảng sau đây. Các mục chính trong mỗi kịch bản kiểm thử bao gồm:

- **TEST SCENARIO:** Tên kịch bản kiểm thử
- **Screen name/Function name:** Tên màn hình hoặc tên chức năng cần kiểm thử
- **Test case code:** Mã số của test case
- **Number of passed test cases (P):** Số lượng test case đã thành công
- **Number of failed test cases (F):** Số lượng test case không thành công
- **Number of test cases under pending (PE):** Số lượng test case đang chờ xử lý
- **Number of unexecuted test cases:** Số lượng test case chưa được thực hiện
- **Total number of test cases:** Tổng số test case cần kiểm thử

#### 6.4.3.1 Cấu trúc của một kịch bản kiểm thử

- Test case code:** Mã của test case, giúp phân biệt các test case khác nhau
- Testing Purpose:** Mục đích của việc kiểm thử
- Steps:** Các bước thực hiện kiểm thử
- Expected outcome:** Kết quả mong đợi sau khi thực hiện các bước kiểm thử
- Browser compatibility testing:** Kiểm tra tính tương thích với các trình duyệt phổ biến
- Test status report:** Báo cáo tình trạng của các lần kiểm thử

#### 6.4.3.2 Ví dụ về một kịch bản kiểm thử trong bảng Excel

TEST SCENARIO *												
Test case code	Testing Purpose	Steps	Expected outcome	Google Chrome	Firefox	Microsoft Edge	Cốc Cốc	Curent	Error code	Note		
<b>TEST STATUS REPORT</b>												
1. Function description: POST /activities/ 2. Test executor: Phan Pham Thi 3. Test execution day: 27/11/2024												
AA_Add_Activity_1	Verify adding a valid activity	1. Create a valid AddActivityRequest with required fields (student_id, type, description). 2. Send a POST request to /activities/ with the valid data.	The activity is added successfully, and the response status is 200 with message "Successfully added the activity."									
AA_Add_Activity_2	Verify error when missing student_id	1. Create an AddActivityRequest without student_id. 2. Send a POST request to /activities/ with the invalid data.	The API returns a 400 BadRequest error with message "Student ID, activity type, and activity description are required."	P	P							

Figure 6.6: Ví dụ về một kịch bản kiểm thử

#### 6.4.3.3 Mục tiêu của các kịch bản kiểm thử

- Kiểm thử chức năng:** Đảm bảo các API hoạt động đúng với các yêu cầu và dữ liệu đầu vào hợp lệ hoặc không hợp lệ.
- Kiểm thử tương thích trình duyệt:** Đảm bảo hệ thống hoạt động ổn định trên các trình duyệt phổ biến.
- Kiểm thử độ tin cậy:** Kiểm tra xem các API có thể xử lý được các tình huống thực tế và đưa ra kết quả chính xác.

#### 6.4.3.4 Kết quả kiểm thử API

##### 6.4.3.4.a Kiểm thử phương thức

- Welcome API Testing**



Testing Purpose	Steps	Expected outcome	Postman
			1st try 2nd try 3rd try
<b>TEST STATUS REPORT</b>			
1. Function description: POST /v1/dashboard/welcome/{studentId}			
2. Test executor: Phan Pham Th			
3. Test execution day: 27/11/2024			
4. Pre-condition: Access link deploy backend: <a href="https://edumind.onrender.com">https://edumind.onrender.com</a> and open testing tool (POSTMAN SWAGGER)			
<i>Verify the retrieval of recent activities for a student</i>			
Retrieve Welcome Message with Valid Student ID	1. Make a GET request to: <code>/v1/dashboard/welcome/{studentId}</code> with a valid student ID in POSTMAN SWAGGER Endpoint: <a href="https://edumind.onrender.com/api/v1/dashboard/welcome/abc123987654">https://edumind.onrender.com/api/v1/dashboard/welcome/abc123987654</a> Method: GET Body: None 2. Click "Send"	1. Response code: 200 2. Response body is not empty 3. Message should be a string 4. Data should contain the course, course_id, and last_accessed <pre>{   "isSuccess": true,   "errCode": 200,   "message": "Successfully fetched the welcome message.",   "messageCode": "SUCCESS",   "data": {     "course": "Data Structures and Algorithms (DSA)",     "course_id": "abc12345-4f39-4714-acb0-6345210",     "last_accessed": "2024-11-07T18:38:15Z"   } }</pre>	P - P - P -
Retrieve Welcome Message with Invalid Student ID	1. Make a GET request to: <code>/v1/dashboard/welcome/{studentId}</code> with a valid student ID in POSTMAN SWAGGER Endpoint: <a href="https://edumind.onrender.com/api/v1/dashboard/welcome/abc123987654">https://edumind.onrender.com/api/v1/dashboard/welcome/abc123987654</a> Method: GET Body: None 2. Click "Send"	1. Response code: 400 2. Response body is not empty 3. Error message should indicate validation failure <pre>{   "isSuccess": false,   "errCode": 400,   "message": "Validation error occurred",   "messageCode": "ERROR",   "data": null }</pre>	P - P - P -

**Figure 6.7: Kiểm thử truy xuất các khóa học gần đây**

Minh họa cách hệ thống truy xuất thông tin về các khóa học gần đây mà sinh viên đã tham gia hoặc có liên quan.

#### • Get Activities API Testing

Testing Purpose	Steps	Expected outcome	Postman	Swagger	Current outcome			
			1st try	2nd try	3rd try	1st try	2nd try	3rd try
<b>TEST STATUS REPORT</b>								
1. Function description: POST /v1/dashboard/activities								
2. Test executor: Phan Pham Th								
3. Test execution day: 27/11/2024								
4. Pre-condition: Access link deploy backend: <a href="https://edumind.onrender.com">https://edumind.onrender.com</a> and open testing tool (POSTMAN SWAGGER)								
<i>Verify the retrieval of student activities with offset</i>								
Retrieve Activities with Default Offset	1. Make a POST request to: <code>/v1/dashboard/activities</code> with the body containing a valid student ID. 2. Ensure the response includes activity data in the data field with an array of activities. Input: - Endpoint: <a href="https://edumind.onrender.com/api/v1/dashboard/activities">https://edumind.onrender.com/api/v1/dashboard/activities</a> - Method: POST - Body: <pre>{   "student_id": "5d5b6ac7-2d7c-4c9d-a5ef-abc123987654",   "limit": 10 }</pre>	* HTTP 200 status code * isSuccess = true * Returns list of activities * Verify activity_id, activity_description, activity_type, activity_date Response should be: <pre>{   "isSuccess": true,   "errCode": 200,   "message": "Successfully fetched the recent activities.",   "messageCode": "SUCCESS",   "data": [     {       "activity_id": "a1bc587-5cc1-4f53-829d-cb6937b29338",       "activity_description": "You have just viewed the course!",       "activity_type": "view_course",       "activity_date": "2024-11-07T16:53:08.635Z"     },     ...   ] }</pre>	P - P - P - P - P - P - P -					
Retrieve Activities with Specific Limit	1. Make a POST request to: <code>/v1/dashboard/activities</code> with the body containing a valid student ID. 2. Ensure the response includes activity data in the data field with an array of activities. Input: - Endpoint: <a href="https://edumind.onrender.com/api/v1/dashboard/activities">https://edumind.onrender.com/api/v1/dashboard/activities</a> - Method: POST - Body: <pre>{   "student_id": "5d5b6ac7-2d7c-4c9d-a5ef-abc123987654",   "limit": 5 }</pre>	* Returns maximum 5 activities * Correct pagination	P - P - P - P -					
Retrieve Activities with Offset	1. Make a POST request to: <code>/v1/dashboard/activities</code> with the body containing a valid student ID. 2. Ensure the response includes activity data in the data field with an array of activities. Input: - Endpoint: <a href="https://edumind.onrender.com/api/v1/dashboard/activities">https://edumind.onrender.com/api/v1/dashboard/activities</a> - Method: POST - Body: <pre>{   "student_id": "5d5b6ac7-2d7c-4c9d-a5ef-abc123987654",   "limit": 5,   "offset": 2 }</pre>	* Returns activities starting from 3rd record * Correct pagination	P - P - P - P -					

**Figure 6.8: Kiểm thử truy xuất các hoạt động gần đây**

Hiển thị kết quả kiểm thử cho việc truy xuất các hoạt động gần đây của sinh viên trong hệ thống.

#### • Courses List API Testing



TEST STATUS REPORT								
1. Function description: POST /courses/ 2. Test executor: Phan Pham Thi 3. Test execution day: 27/11/2024 4. Pre-Condition: Access link deploy backend: <a href="https://edumind.onrender.com">https://edumind.onrender.com</a> and open testing tool (POSTMAN/SWAGGER) <i>Verify retrieval of course list for a student</i>								
Previous Condition								
Pagination	1. Make a POST request to /v1/courses/ with the body containing a valid student ID. 2. Ensure the response includes courses data in the data field with an array of courses' information Input: - Endpoint: <a href="https://edumind.onrender.com/api/v1/courses/">https://edumind.onrender.com/api/v1/courses/</a> - Method: POST - Body: ( "student_id": "5d3bdaac7-2d7c-4cd4-a5ef-abcl23987654" )	* HTTP 200 status code * isSuccess = true * Returns list of courses * Verify course details: - id - name - start_date, end_date - student_list - learning_outcomes - professor_details - status - image - percentage_complete - last_accessed	P +	P +	P +	P +	P +	P +
Retrieve Courses with Valid Custom Page Size	1. Make a POST request to /v1/courses/ with the body containing a valid student ID and valid custom page size 2. Ensure the response includes courses data in the data field with an array of courses' information Input: - Endpoint: <a href="https://edumind.onrender.com/api/v1/courses/">https://edumind.onrender.com/api/v1/courses/</a> - Method: POST - Body: ( "student_id": "5d3bdaac7-2d7c-4cd4-a5ef-abcl23987654", "page_size": "100" )	* Returns maximum 100 courses * Correct pagination metadata	P +	P +	P +	P +	P +	P +
Retrieve Courses with Invalid Custom Page Size	1. Make a POST request to /v1/courses/ with the body containing a valid student ID and invalid custom page size 2. Ensure the response includes courses data in the data field with an array of courses' information Input: - Endpoint: <a href="https://edumind.onrender.com/api/v1/courses/">https://edumind.onrender.com/api/v1/courses/</a> - Method: POST - Body: ( "student_id": "5d3bdaac7-2d7c-4cd4-a5ef-abcl23987654", "page_size": "-100" )	* Returns maximum 100 courses * Correct pagination metadata	F -	F -	F -	F -	F -	F -
Retrieve Courses with Large Offset	1. Make a POST request to /v1/courses/ with the body containing a valid student ID and valid custom page size 2. Ensure the response includes courses data in the data field with an array of courses' information Input: - Endpoint: <a href="https://edumind.onrender.com/api/v1/courses/">https://edumind.onrender.com/api/v1/courses/</a> - Method: POST - Body: ( )	* Returns appropriate response * Handles large offset scenarios	F -	F -	F -	F -	F -	F -

**Figure 6.9: Kiểm thử truy xuất danh sách khóa học của sinh viên**

Thể hiện chi tiết kết quả khi lấy danh sách các khóa học mà sinh viên đã đăng ký hoặc tham gia.

#### • Get Course for Student API Testing

TEST STATUS REPORT								
1. Function description: GET /courses/{courseId}/students/{studentId}/ 2. Test executor: Phan Pham Thi 3. Test execution day: 27/11/2024 4. Pre-Condition: Access link deploy backend: <a href="https://edumind.onrender.com">https://edumind.onrender.com</a> and open testing tool (POSTMAN/SWAGGER) <i>Verify retrieval of specific course details for a student</i>								
Previous Condition								
Retrieve Course Details with Valid IDs	* Valid courseId * Valid studentId	* HTTP 200 status code * isSuccess = true * Verify course details: - courseId - student_id - completed_lessons - time_spent - assignments_done - lessons_details: - id - title_description - lesson_type - bookmark_status - order - status - exercises_within_lessons - documents_associated_with_lessons	P +	P +	P +	P +	P +	P +
Retrieve Course with Mismatched Course and Student ID	* Valid courseId * Different studentId	* Appropriate error handling * Consistent error response Response should be: { "success": false, "errorCode": 400, "message": "Student is not enrolled in this course.", "messageCode": "ERROR", "data": null	P +	P +	P +	P +	P +	P +
Retrieve Course with Invalid IDs	* Invalid courseId * Invalid studentId	* HTTP 404 status code * Validation error details	F -	F -	F -	F -	F -	F -

**Figure 6.10: Kiểm thử truy xuất chi tiết một khóa học**

Kết quả trả về chi tiết thông tin của một khóa học cụ thể trong hệ thống.

- **Recommended Courses List API Testing**

Testing Purpose	Steps	Expected outcome	Postman	Swagger	Current				
TEST STATUS REPORT			1st try	2nd try	3rd try	1st try	2nd try	3rd try	outcome
<i>1. Function description: GET '/courseId/students/studentId/lessons_recommendation'</i>									
<i>2. Test executor : Phan Pham Thi</i>									
<i>3. Test execution day: 27/11/2024</i>									
<i>4. Pre-Condition: Access link deploy backend: https://edumind.onrender.com/ and open testing tool (POSTMAN/SWAGGER)</i>									
<b>Verify retrieval of recommended course list for a student</b>									
<i>Preview Condition</i>									
Retrieve Courses with Default Pagination	Input: - student_id: Valid student UUID - course_id: Valid courses UUID Endpoint: <a href="https://edumind.onrender.com/api/v1/courses/abc12345-6789-4321-fedc-ba9876543210/students/d5b6ac7-247c-4e9d-a5ef-abc123987654/lessons_recommendation/">https://edumind.onrender.com/api/v1/courses/abc12345-6789-4321-fedc-ba9876543210/students/d5b6ac7-247c-4e9d-a5ef-abc123987654/lessons_recommendation/</a>	* HTTP 200 status code * isSuccess = true * Returns list of courses * Verify course details: - course_id - course_name: str - lesson_id: UUID - bookmark: bool - status: StatusType - title: str - description: str - order: int	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓
Retrieve Courses with Invalid Course ID	Input: - student_id: Valid student UUID - course_id: Invalid courses UUID	* HTTP 500 status code * Validation error details { "isSuccess": false, "errorCode": 404, "message": "Learning path not found for the given student and course.", "messageCode": "ERROR", "data": null }	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓
Retrieve Courses with Invalid Student ID	Input: - student_id: Invalid student UUID - offset: 100 - page_size: 10	* HTTP 500 status code * Validation error details	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓

**Figure 6.11: Kiểm thử truy xuất danh sách khóa học được đề xuất**

Truy xuất danh sách các khóa học được hệ thống đề xuất dựa trên sở thích hoặc lịch sử học tập của sinh viên.

- **Add Activity API Testing**

Testing Purpose	Steps	Expected outcome	Postman	Swagger	Current				
TEST STATUS REPORT			1st try	2nd try	3rd try	1st try	2nd try	3rd try	outcome
<i>1. Function description: POST /activities/</i>									
<i>2. Test executor : Phan Pham Thi</i>									
<i>3. Test execution day: 27/11/2024</i>									
<b>Verify adding a valid activity</b>									
Verify adding a valid activity	1. Create a valid AddActivityRequest with required fields (student_id, type, description). 2. Send a POST request to /activities/ with the valid data.	The activity is added successfully, and the response status is 200 with message "Successfully added the activity."	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓
Verify error when missing student_id	1. Create an AddActivityRequest without student_id. 2. Send a POST request to /activities/ with the invalid data.	The API returns a 400 BadRequest error with message "Student ID, activity type, and activity description are required." { "isSuccess": false, "errorCode": 400, "message": "Validation error occurred", "messageCode": "ERROR", "data": null }	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓
Verify error when missing type	1. Create an AddActivityRequest without type. 2. Send a POST request to /activities/ with the invalid data.	The API returns a 400 BadRequest error with message "Student ID, activity type, and activity description are required." { "isSuccess": false, "errorCode": 400, "message": "Validation error occurred", "messageCode": "ERROR", "data": null }	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓
Verify error when missing description	1. Create an AddActivityRequest without description. 2. Send a POST request to /activities/ with the invalid data	The API returns a 400 BadRequest error with message "Student ID, activity type, and activity description are required." { "isSuccess": false, "errorCode": 400, "message": "Validation error occurred", "messageCode": "ERROR", "data": null }	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓
Verify error when activity fails to be added to the database	1. Create a valid AddActivityRequest. 2. Mock the activities_controller.activities_repository.create method to return None. 3. Send a POST request to /activities/.	The API returns a 404 NotFound error with message "Failed to add activity."	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓	P ✓

**Figure 6.12: Kiểm thử thêm hoạt động gần đây của sinh viên**

Min họa kiểm thử chức năng ghi nhận một hoạt động mới vào danh sách các hoạt động gần đây của sinh viên.

- **Recommend Lesson API Testing**



Testing Purpose	Steps	Expected outcome	Postman	Swagger	Current outcome
<b>1. Function description: GET /v1/recommendedLessons/{recommendedLessonId}</b>					
<b>2. Test executor: Applitools Testbar</b>					
<b>3. Test execution day: 27/11/2023</b>					
<b>Verify the functionality of retrieving a lesson message for a student</b>					
Retrieve Recommended Lesson with Valid ID	1. Provide a valid recommendedLessonId. EndPoint: <a href="https://edumind.onrender.com/api/v1/recommendedLessons/4e4ff5fcf1f8a4133f987b-654321ffedcba">https://edumind.onrender.com/api/v1/recommendedLessons/4e4ff5fcf1f8a4133f987b-654321ffedcba</a> Method: GET Body: None 2. Send request	<pre>* HTTP 200 status code * isSuccess = true * Response data includes: lesson_id, name, learning_outcomes, description, progress, status, recommended_content, explain_modules {     "isSuccess": true,     "errorCode": 0,     "message": "Successfully fetched the recommended lesson.",     "data": {         "lesson_id": "4e4ff5fcf1f8a4133f987b-654321ffedcba",         "name": "Introduction to Data Structures",         "description": "Understand fundamental concepts of Data Structures",         "learning_outcomes": "Overview of fundamental data structures and their applications",         "status": "Completed"     } }</pre>	P ~	P ~	P ~
Retrieve Lesson with Invalid ID	1. Provide an invalid recommendedLessonId. EndPoint: <a href="https://edumind.onrender.com/api/v1/recommendedLessons/4e4ff5fcf1f8a4133f987b-654321ffedcba">https://edumind.onrender.com/api/v1/recommendedLessons/4e4ff5fcf1f8a4133f987b-654321ffedcba</a> Method: GET Body: None 2. Send request	<pre>* isSuccess = false, * errorCode = 404, * message = "Recommended Lesson not found for the given ID." * data = null</pre>	P ~	P ~	P ~
Retrieve Lesson Without Required Relations	1. Provide a valid recommendedLessonId but the lesson has no associated modules or relations. EndPoint: <a href="https://edumind.onrender.com/api/v1/recommendedLessons/d0a983c7-e6f9-7a8d-cba9-21fedcba45f7">https://edumind.onrender.com/api/v1/recommendedLessons/d0a983c7-e6f9-7a8d-cba9-21fedcba45f7</a> Method: GET Body: None 2. Send request	<pre>* HTTP 200 status code * isSuccess = true * Response includes lesson details but modules is an empty array {     "isSuccess": true,     "errorCode": 0,     "message": "Successfully fetched the recommended lesson.",     "data": {         "lesson_id": "d0a983c7-e6f9-7a8d-cba9-21fedcba45f7",         "name": "Dynamic Programming",         "description": "Learn array indexing",         "modules": []     } }</pre>	P ~	P ~	P ~

**Figure 6.13: Kiểm thử truy xuất chi tiết một lesson được đề xuất**

Thể hiện kết quả kiểm thử khi truy xuất chi tiết về một bài học được hệ thống đề xuất.

- **Quizzes List API Testing**

Testing Purpose	Steps	Expected outcome	Postman	Swagger	Current outcome
<b>1. Function description: GET /v1/modules/{moduleId}/quizzes</b>					
<b>2. Test executor: Applitools Testbar</b>					
<b>3. Test execution day: 27/11/2023</b>					
<b>Verify the functionality of retrieving quizzes by module ID</b>					
Retrieves a Module with Quizzes Using Valid Module ID	1. Provide a valid moduleId with associated quizzes. EndPoint: <a href="https://edumind.onrender.com/api/v1/modules/b7d8c920-3a4b-4b6e-a8c1-84c96c90e4f3/quizzes">https://edumind.onrender.com/api/v1/modules/b7d8c920-3a4b-4b6e-a8c1-84c96c90e4f3/quizzes</a> Method: GET Body: None 2. Send request	<pre>* HTTP 200 status code * isSuccess = true * Response data includes: module_id, title, objectives, quizzes (with details: id, name, status, difficulty, score) {     "isSuccess": true,     "errorCode": 0,     "message": "Successfully fetched the module and quizzes.",     "data": {         "module_id": "b7d8c920-3a4b-4b6e-a8c1-84c96c90e4f3",         "title": "Learning Array",         "objectives": [             "Learn array indexing",             "Apply sorting algorithms on arrays"         ],         "quizzes": [             {                 "id": "066d8c9e-9e0b-447c-a8f1-12cbca54321ef",                 "name": "Quiz 1"             }         ]     } }</pre>	P ~	P ~	P ~
Retrieve Module Without Quizzes Using Valid Module ID	1. Provide a valid moduleId with no associated quizzes. EndPoint: <a href="https://edumind.onrender.com/api/v1/modules/b7d8c920-3a4b-4b6e-a8c1-84c96c90e4f3/quizzes">https://edumind.onrender.com/api/v1/modules/b7d8c920-3a4b-4b6e-a8c1-84c96c90e4f3/quizzes</a> Method: GET Body: None 2. Send request	<pre>* HTTP 200 status code * isSuccess = true * Response data includes valid module details but quizzes is an empty array {     "isSuccess": true,     "errorCode": 0,     "message": "Successfully fetched the module and quizzes.",     "data": {         "module_id": "b7d8c920-3a4b-4b6e-a8c1-84c96c90e4f3",         "title": "Learning Array",         "objectives": [             "Learn array indexing",             "Apply sorting algorithms on arrays"         ],         "quizzes": []     } }</pre>	P ~	P ~	P ~
Retrieve Module with Invalid Module ID	1. Provide an invalid moduleId (non-existent in the database) EndPoint: <a href="https://edumind.onrender.com/api/v1/modules/c5e4d09e-9e0b-447c-a8f1-12cbca54321ef/quizzes">https://edumind.onrender.com/api/v1/modules/c5e4d09e-9e0b-447c-a8f1-12cbca54321ef/quizzes</a> Method: GET Body: None 2. Send request	<pre>* HTTP 200 status code * isSuccess = false * errorCode = 404 * message = "Module not found for the given ID." * data = null</pre>	P ~	P ~	P ~
Retrieve Module When moduleId Is Missing	1. Do not include a moduleId in the request URL 2. Send a GET request to /v1/modules/quizzes	For moduleId: Required field is not provided.	P ~	P ~	P ~

**Figure 6.14: Kiểm thử truy xuất danh sách bài tập quizzes từ module**

Hiển thị kết quả kiểm thử danh sách bài tập dạng quiz từ một module cụ thể.

- **Quiz Detail API Testing**



Testing Purpose	Steps	Expected outcome	Postman	Swagger	Current outcome			
			1st try	2nd try	3rd try	1st try	2nd try	3rd try
<b>TEST STATUS REPORT</b>								
1. Function description: GET /v1/modules/{moduleId}/quizzes/{quizId}								
2. Test executor: Python Testing Tools API								
<b>Verify the functionality of retrieving a quiz detail message for a student</b>								
Retrieve Quiz Details with Valid moduleId and quizId	1. Provide a valid moduleId and quizId. Endpoint: https://edumind.comendar.com/api/v1/modules/b7d0-020-3ab-4bde-a8c1-84a96-90e4f/quizzes/0188535f-fbd4-4fb4-b745-553d997a6fba Method: GET. Body: None. 2. Send request	* HTTP 200 status code * isSuccess = true * Response data includes: id, name, status, difficulty, score, question/question, options, explanation, user_choice, correct_answer, image	P -	P -	P -	P -	P -	P -
Retrieve Quiz with Invalid moduleId	1. Provide a non-existent or invalid moduleId. Endpoint: https://edumind.comendar.com/api/v1/modules/b7d0-020-3ab-4bde-a8c1-84a96-90e4f/quizzes/0188535f-fbd4-4fb4-b745-553d997a6fba Method: GET. Body: None. 2. Send request	* HTTP 404 status code. * isSuccess = false * Error message: "Module not found for the given ID."	P -	P +	P -	P -	P -	P -

**Figure 6.15: Kiểm thử truy xuất chi tiết một bài tập quiz**

Kết quả kiểm thử truy xuất thông tin chi tiết về một bài tập quiz trong danh sách.

### • Clear Answer API Testing

Testing Purpose	Steps	Expected outcome	Postman	Swagger	Current outcome	Error code	Note	
			1st try	2nd try	3rd try	1st try	2nd try	3rd try
<b>TEST STATUS REPORT</b>								
1. Function description: PUT /v1/modules/{moduleId}/quizzes/{quizId}/submit								
2. Test executor: Ngopyin Testing Tools API								
3. Test environment: Local								
<b>Verify the functionality of clearing answer in a quiz</b>								
Successful Reset of Quiz Answers	1. Provide a valid moduleId and quizId. Endpoint: https://edumind.comendar.com/api/v1/modules/b7d0-020-3ab-4bde-a8c1-84a96-90e4f/quizzes/0188535f-fbd4-4fb4-b745-553d997a6fba/clear Method: PUT. Body: None. 2. Send request	* HTTP 200 status code * isSuccess = true * Quiz status is reset to in_progress. * Score is reset to 0.	P -	P -	P -	P -	P -	P -
Quiz Not Found for Given Quiz ID	1. Provide a valid moduleId but an invalid quizId. Endpoint: https://edumind.comendar.com/api/v1/modules/b7d0-020-3ab-4bde-a8c1-84a96-90e4f/quizzes/0188535f-fbd4-4fb4-b745-553d997a6fba/clear Method: PUT. Body: None. 2. Send request	* HTTP 404 status code * isSuccess = false * Error message: "Quiz not found for the given ID".	P -	P -	P -	P -	P -	P -
Quiz with No Existing Answers	1. Provide a valid moduleId and quizId. 2. Ensure quiz has no submitted answers. Endpoint: https://edumind.comendar.com/api/v1/modules/b7d0-020-3ab-4bde-a8c1-84a96-90e4f/quizzes/0188535f-fbd4-4fb4-b745-553d997a6fba/clear Method: PUT. Body: None. 2. Send request.	* HTTP 200 status code * isSuccess = true * Quiz remains in in_progress with score as 0.	P -	P -	P -	P -	P -	P -

**Figure 6.16: Kiểm thử bắt đầu làm lại một bài tập quiz**

Thể hiện kết quả thử nghiệm chức năng cho phép sinh viên bắt đầu làm lại một bài tập quiz đã làm trước đó.

### • Submit Answer API Testing



Testing Purpose	Steps	Expected outcome	Postman			Swagger			Current outcome			
			1st try	2nd try	3rd try	1st try	2nd try	3rd try				
<b>TEST STATUS REPORT</b>												
1. Function description: POST /v1/modules/{moduleId}/quizzes/{quizId}/submit 2. Test executor: Xpoiler Testing Tools And 3. Test execution date: 17/11/2024												
<b>Verify the functionality of submitting a quiz</b>												
Successful Submission with All Correct Answers	1. Provide a valid moduleId and quizId. 2. Submit answers that match all correct options. moduleId: b7a8e20-3a4b-4b6e-a513-34c9e90e4fB quizId: 0188335f-fd44-4fb0-b743-553d997a6fea - Endpoints: https://edumind.com/api/v1/modules/b7a8e20-3a4b-4b6e-a513-34c9e90e4fB/quizzes/0188335f-fd44-4fb0-b743-553d997a6fea - Method: PUT - Body: ( "answers": [0,3] )	* HTTP 200 status code * isSuccess = true * Quiz answers match the number of questions. * Status updated to completed. * Full score (100%). { "isSuccess": true, "errorCode": 0, "message": "Quiz answers submitted successfully.", "messageCode": "SUCCESS", "data": { "quiz_id": "0188335f-fd44-4fb0-b743-553d997a6fea", "total_questions": 2, "correct_answers": 2, "score": 100, "results": [ { "id": 1, "score": 100 } ] } }	P	P	P	P	P	P	P			
Successful Submission with Partial Correct Answers	1. Provide a valid moduleId and quizId. 2. Submit partially correct answers. moduleId: b7a8e20-3a4b-4b6e-a513-34c9e90e4fB quizId: 0188335f-fd44-4fb0-b743-553d997a6fea - Endpoints: https://edumind.com/api/v1/modules/b7a8e20-3a4b-4b6e-a513-34c9e90e4fB/quizzes/0188335f-fd44-4fb0-b743-553d997a6fea - Method: PUT - Body: ( "answers": [0,3] )	* HTTP 200 status code * isSuccess = true * Status updated to completed. * Partial score calculated correctly. { "isSuccess": true, "errorCode": 0, "message": "Quiz answers submitted successfully.", "messageCode": "SUCCESS", "data": { "quiz_id": "0188335f-fd44-4fb0-b743-553d997a6fea", "total_questions": 2, "correct_answers": 1, "score": 50, "results": [ { "id": 1, "score": 50 } ] } }	P	P	P	P	P	P	P			
Submission with All Incorrect Answers	1. Provide a valid moduleId and quizId. 2. Submit answers where all are incorrect. moduleId: b7a8e20-3a4b-4b6e-a513-34c9e90e4fB quizId: 0188335f-fd44-4fb0-b743-553d997a6fea - Endpoints: https://edumind.com/api/v1/modules/b7a8e20-3a4b-4b6e-a513-34c9e90e4fB/quizzes/0188335f-fd44-4fb0-b743-553d997a6fea - Method: PUT - Body: ( "answers": [3,3] )	* HTTP 200 status code * isSuccess = true * Status updated to completed. * Score is 0. { "isSuccess": true, "errorCode": 0, "message": "Quiz answers submitted successfully.", "messageCode": "SUCCESS", "data": { "quiz_id": "0188335f-fd44-4fb0-b743-553d997a6fea", "total_questions": 2, "correct_answers": 0, "score": 0, "results": [ { "id": 1, "score": 0 } ] } }	P	P	P	P	P	P	P			

**Figure 6.17: Kiểm thử submit một bài tập quiz**

Kiểm tra chức năng nộp bài làm cho một bài tập quiz thông qua hệ thống.

#### • Document API Testing

Testing Purpose	Steps	Expected outcome	Postman			Swagger			Current outcome			
			1st try	2nd try	3rd try	1st try	2nd try	3rd try				
<b>TEST STATUS REPORT</b>												
1. Function description: POST /v1/modules/{moduleId}/documents 2. Test executor: Xpoiler Testing Tools And 3. Test execution date: 17/11/2024												
<b>Verify the functionality of retrieving a document from module for a student</b>												
Retrieve Document Details with Valid moduleId	1. Provide a valid moduleId. Endpoint: https://edumind.com/api/v1/modules/b7a8e20-3a4b-4b6e-a513-34c9e90e4f/documents Method: GET Body: None 2. Send request	* HTTP 200 status code * isSuccess = true * Response data includes: id, name, theoryContent/title, examples, descriptions, practicalContent, steps, title, commonError, references, link, title, description { "isSuccess": true, "errorCode": 0, "message": "Successfully fetched the document.", "messageCode": "SUCCESS", "data": { "id": "141-24-1a1f4f2-3d7-0173e8bb5e5", "name": "Milestone Array", "theoryContent": [ { "title": "Understanding Arrays", "prerequisites": [ { "id": "141-24-1a1f4f2-3d7-0173e8bb5e5", "name": "Basic Data Types" } ] } ] } }	P	P	P	P	P	P	P			
Retrieve Document with Invalid moduleId	1. Provide a non-existent or invalid moduleId. Endpoint: https://edumind.com/api/v1/modules/b7a8e20-3a4b-4b6e-a513-34c9e90e4f/documents Method: GET Body: None 2. Send request	* HTTP 404 status code. * isSuccess = false * Error message: "Document not found for the given ID."	P	P	P	P	P	P	P			
Retrieve Document with Empty or Missing Content	1. Provide a document where the content field is empty or missing. Endpoint: https://edumind.com/api/v1/modules/b7a8e20-3a4b-4b6e-a513-34c9e90e4f/documents Method: GET Body: None 2. Send request	* HTTP 400 status code. * isSuccess = false * Error message: "Document content is missing or empty."	P	T	T	T	T	T	T			
Retrieve Quiz with Invalid Questions Format	1. Simulate invalid content data (e.g., missing required keys like name, references, ... ) in the database. Endpoint: https://edumind.com/api/v1/modules/b7a8e20-3a4b-4b6e-a513-34c9e90e4f/documents Method: GET Body: None Content must have field title = name 2. Send request	* HTTP 500 status code. * isSuccess = false * Error message: "Invalid data format in question: <missing_key>."	T	T	T	T	T	T	T			

**Figure 6.18: Kiểm thử truy xuất document từ module**

Minh họa việc truy xuất tệp tài liệu từ một module khóa học.

#### 6.4.3.4.b Kết quả tổng hợp của các testcase

SYNTHESIS OF RESULTS										
ID	Screen name/Testcase name	Tests passed (%)	Tests failed (%)	Tests pending (%)	Number of unexpected test cases	Number of executed test cases	Percentage of test cases passed (%)	Percentage of test cases pending (%)	Percentage of test cases failed (%)	Percentage of test cases executed (%Cover)
SIG1	Welcome API Testing	2	0	0	0	2	100%	0%	0%	100%
SIG2	Get Activities Testing	4	0	0	0	4	100%	0%	0%	100%
SIG3	Courses List API Testing	3	2	0	0	5	60%	0%	40%	100%
SIG4	Create Course for Student API Testing	2	0	0	0	2	100%	0%	0%	100%
SIG5	Get Activities List API Testing	3	0	0	0	3	100%	0%	0%	100%
SIG6	Add Activity API Testing	7	0	0	0	7	100%	0%	0%	100%
SIG7	Get Courses List API Testing	3	0	0	0	3	100%	0%	0%	100%
SIG8	Get Quizzes List API Testing	4	0	0	0	4	100%	0%	0%	100%
SIG9	Quiz Detail API Testing	4	0	0	0	4	100%	0%	0%	100%
SIG10	Document API Testing	2	2	0	0	6	33%	0%	50%	100%
SIG11	Document Detail API Testing	3	0	0	0	3	100%	0%	0%	100%
SIG12	Clear Answer API Testing	3	0	0	0	3	100%	0%	0%	100%
SIG13	Clear Answer API Testing	Total	42	6	0	48	88%	0%	13%	100%

**Figure 6.19: Kết quả tổng hợp của các testcase**



- Kết quả tổng quan:
  - **Số lượng test case đã thực thi:** Toàn bộ 48 test case đã được thực thi, đạt tỷ lệ bao phủ kiểm thử là 100%.
  - **Số lượng test case đã vượt qua:** Có 42 test case thành công, chiếm 88% tổng số test case.
  - **Số lượng test case thất bại:** Có 6 test case thất bại, chiếm 13% tổng số test case.
  - **Số lượng test case đang chờ xử lý:** Không có test case nào đang chờ xử lý (0%).
- Phân tích chi tiết theo từng chức năng
  - **Welcome API Testing, Get Activities Testing, Recommended Courses List API Testing, Add Activity API Testing, Recommend Lesson API Testing, Quizzes List API Testing, Quiz Detail API Testing, Clear Answer API Testing:** Các chức năng này đều đạt tỷ lệ thành công 100%, không có lỗi xảy ra trong quá trình kiểm thử.
  - **Courses List API Testing:** Tỷ lệ thành công đạt 60% với 3 test case vượt qua và 2 test case thất bại.
  - **Get Course for Student API:** Tỷ lệ thành công đạt 67%, với 2 test case vượt qua và 1 test case thất bại.
  - **Document API Testing:** Có tỷ lệ thành công thấp nhất (50%), với 2 test case vượt qua và 2 test case thất bại.
  - **Submit Answer API Testing:** Đạt tỷ lệ thành công 83%, với 5 test case vượt qua và 1 test case thất bại.

#### **6.4.3.4.c Hạn chế**

- Do nhóm chưa có nhiều kinh nghiệm trong việc thực hiện unit testing.
- Quá trình kiểm thử chủ yếu tập trung vào kiểm tra chức năng API mà chưa thực hiện đầy đủ các kiểm thử liên quan đến giao diện người dùng (UI testing) hoặc hiệu năng hệ thống (performance testing).
- Một số API vẫn còn thất bại, đặc biệt là Courses List API Testing và Document API Testing, cần được kiểm tra và sửa lỗi.



## 6.5 Thủ nghiệm hệ thống với các mô hình ngôn ngữ hiện có

Mục đích của cuộc thử nghiệm đó là đánh giá chất lượng của lời giải thích của từng mô hình ngôn ngữ hồn dành cho người học. Mô hình tiến hành thử nghiệm:

- GPT 4o
- GPT 4o Mini
- Google Gemini Pro 1.5

Dữ liệu đầu vào cho mỗi mô hình là một yêu cầu từ người dùng về việc cung cấp các tài nguyên học tập hữu ích cho việc phát triển kỹ năng. Dưới đây là kết quả trả về từ mỗi mô hình:

- GPT 4o mini:

```
1      GPT 4o mini:  
2      Response body  
3      {  
4          "final_response": [  
5              {  
6                  "id": "b0f66328-862d-4e4d-9f5a-2bdee93490b8",  
7                  "code": "COD001",  
8                  "description": "A classic beginner coding challenge.",  
9                  "explanation": "This coding challenge is a great  
10                     starting point for you as it will help you practice  
11                     fundamental programming skills. Engaging with such  
12                     challenges will enhance your problem-solving  
13                     abilities, which are essential for backend  
14                     development."  
15              },  
16              {  
17                  "id": "d3f325a9-8103-4a7b-bb26-7de45b607f9c",  
18                  "code": "COD002",  
19                  "description": "A simple exercise to reverse a given  
20                     string.",  
21                  "explanation": "This exercise will allow you to work on  
22                     string manipulation, a critical skill in backend  
23                     development. By mastering this concept, you will be  
24                     better prepared to handle data processing tasks that  
25                     are common in backend applications."  
26              },  
27              {  
28                  "id": "59427089-44c4-4916-8886-52a9f6f488b1",  
29                  "code": "CRS003",  
30              }  
31          }  
32      }  
33  }
```



```
20         "description": "A comprehensive introduction to web
21             development.",
22         "explanation": "This course will provide you with a
23             solid foundation in web development, covering
24             essential concepts that are crucial for backend
25             development. Understanding how the web works will
26             empower you to create effective backend solutions."
27     },
28     {
29         "id": "1144a6d7-a357-4e4c-b4f5-315afde72a67",
30         "code": "PA001",
31         "description": "A project to build a calculator using
32             Python.",
33         "explanation": "Working on this project will give you
34             hands-on experience with Python, a popular backend
35             language. It will help you apply your knowledge in a
36             practical setting, enhancing your coding skills and
37             preparing you for real-world backend tasks."
38     },
39     {
40         "id": "7a32d80f-5313-425f-9bf2-c271084384d8",
41         "code": "PA003",
42         "description": "A project to create and deploy a
43             personal portfolio website.",
44         "explanation": "This project will not only allow you to
45             practice your web development skills but also give
46             you experience in deploying a web application.
47             Understanding deployment is vital for backend
48             developers, as it will help you manage and maintain
49             applications effectively."
50     }
51 ]
52 }
```

- GPT 4o

```
1  {
2      GPT 4o:
3      Response body
4      "final_response": [
5          {
6              "id": "59427089-44c4-4916-8886-52a9f6f488b1",
7              "code": "CRS003",
8              "description": "A comprehensive introduction to web
```



```
    development.",
9   "explanation": "Since you're starting your journey towards
  backend development, having a solid understanding of web
  development is essential. This course will provide you
  with the foundational knowledge necessary to understand
  how web applications are structured, which is crucial
  before diving into backend specifics."
10  },
11  {
12    "id": "CRS001",
13    "code": "6e906d7e-8f93-4dd8-8890-fa4b42cfcc8c",
14    "description": "A beginner-friendly course to learn Python
      programming.",
15    "explanation": "Python is a popular language for backend
      development due to its simplicity and effectiveness. This
      course is perfect for you as it will introduce you to
      programming concepts using Python, which is often used
      for backend tasks such as server scripting and data
      management."
16  },
17  {
18    "id": "PA003",
19    "code": "7a32d80f-5313-425f-9bf2-c271084384d8",
20    "description": "A project to create and deploy a personal
      portfolio website.",
21    "explanation": "Engaging in a project like this will give
      you hands-on experience with web hosting and deployment,
      skills that are vital for a backend developer. You'll
      learn how to manage servers and deploy applications,
      which are key components of backend development."
22  },
23  {
24    "id": "LES002",
25    "code": "1bb2500f-4275-4672-9089-fd9b382809a6",
26    "description": "A lesson focused on understanding binary
      search.",
27    "explanation": "Understanding algorithms is a critical skill
      for backend development, especially for optimizing data
      retrieval processes. This lesson will help you grasp the
      concept of binary search, enhancing your problem-solving
      skills and efficiency in handling data."
28  },
29  {
30    "id": "PA001",
31    "code": "1144a6d7-a357-4e4c-b4f5-315afde72a67",
32    "description": "A project to build a calculator using Python
```



```
    .",
33     "explanation": "This project will reinforce your Python
34         programming skills by applying them to a practical task.
35         It will help you understand how to implement logic and
36         functionality , which is essential for developing backend
37         applications ."
38     }
39   ]
40 }
```

- Gemini Pro 1.5

```
1  {
2    Gemini Pro 1.5:
3    Response body
4    "final_response": [
5      {
6        "id": "b0f66328-862d-4e4d-9f5a-2bdee93490b8",
7        "code": "COD001",
8        "description": "A classic beginner coding challenge.",
9        "explanation": "This coding challenge will help you build
10          foundational programming skills. It's a great way to
11          start practicing problem-solving , a crucial skill for any
12          programmer, especially in backend development where you'
13          ll constantly be tackling new challenges."
14      },
15      {
16        "id": "d3f325a9-8103-4a7b-bb26-7de45b607f9c",
17        "code": "COD002",
18        "description": "A simple exercise to reverse a given string .
19        ",
20        "explanation": "This exercise focuses on string manipulation
21          , a core skill for backend development. Practicing this
22          will improve your ability to process and handle data
23          effectively , a common task in backend applications."
24      },
25      {
26        "id": "CRS001",
27        "code": "CRS001",
28        "description": "A beginner-friendly course to learn Python
29          programming .",
30        "explanation": "Since you're just starting , this course
31          provides a foundational understanding of Python , a widely
32          -used language for backend development. Learning the
```



```
        basics will lay a solid groundwork for more advanced
        concepts and projects later on."
22    },
23    {
24        "id": "LES001",
25        "code": "LES001",
26        "description": "An in-depth lesson on Python lists.",
27        "explanation": "This lesson will build your proficiency in
        working with Python lists, a fundamental data structure.
        Understanding list manipulation is essential for many
        backend programming tasks, improving your overall coding
        efficiency."
28    },
29    {
30        "id": "PA001",
31        "code": "PA001",
32        "description": "A project to build a calculator using Python
        .",
33        "explanation": "This project allows you to apply what you
        learn in a practical way. Building a calculator will
        solidify your understanding of fundamental programming
        concepts like user input, arithmetic operations, and
        function use within the context of a complete application
        ."
34    }
35 ]
36 }
```

Đánh giá về lời giải thích các tài nguyên:

- GPT-4o-mini: Các lời giải thích cho mỗi tài nguyên đều rõ ràng, nhấn mạnh lợi ích mà người học sẽ nhận được, chẳng hạn như việc nâng cao kỹ năng giải quyết vấn đề hoặc làm quen với các công cụ phát triển backend. Tuy nhiên, các lời giải thích có phần chung chung và ít tập trung vào ứng dụng thực tế của các kỹ năng học được.
- GPT-4o: Cung cấp những lời giải thích chi tiết hơn về lý do tại sao mỗi tài nguyên lại hữu ích cho người học, ví dụ như việc hiểu cách hoạt động của web và tầm quan trọng của các kiến thức về máy chủ và triển khai. Những lời giải thích này giúp người học hiểu rõ hơn về mục đích và ứng dụng của các tài nguyên trong việc phát triển kỹ năng backend.
- Gemini Pro: Cũng cung cấp những lời giải thích chi tiết, nhưng có phần thiên về việc khuyến khích người học tham gia vào các bài học thực tế. Mô hình này đưa ra những lý



do cụ thể như việc học cách xử lý chuỗi (COD002) hay hiểu cách danh sách hoạt động trong Python (LES001), từ đó giúp người học nâng cao kỹ năng giải quyết vấn đề và lập trình thực tế.

Về chất lượng và chi tiết giải thích:

- GPT-4o-mini: Lời giải thích trong GPT-4o-mini thường ngắn gọn và tập trung vào lý do chung chung tại sao một tài nguyên lại hữu ích, nhưng thiếu sâu sắc về cách tài nguyên đó giúp người học phát triển các kỹ năng cụ thể.
- GPT-4o: Các lời giải thích chi tiết hơn và cung cấp cái nhìn rõ ràng về các ứng dụng thực tế của từng tài nguyên. Chúng giúp người học nhận thức được lý do tại sao mỗi tài nguyên lại quan trọng trong hành trình học tập của họ.
- Gemini Pro: Cung cấp các lời giải thích rõ ràng, tuy nhiên có xu hướng tập trung vào các kỹ năng cụ thể mà người học sẽ phát triển thông qua các bài học và dự án. Gemini Pro khuyến khích người học áp dụng ngay các kỹ năng vào thực tế thông qua các dự án và bài học.

**Kết luận:** Tất cả ba mô hình GPT-4o-mini, GPT-4o và Gemini Pro đều cung cấp các tài nguyên học tập có giá trị, nhưng mức độ chi tiết và sự rõ ràng trong lời giải thích có sự khác biệt. GPT-4o là mô hình có giải thích chi tiết và nhấn mạnh đến tầm quan trọng của việc hiểu sâu về lý thuyết, trong khi Gemini Pro có xu hướng tập trung vào thực hành nhiều hơn. GPT-4o-mini, mặc dù cung cấp các tài nguyên tương tự, nhưng các lời giải thích có phần đơn giản và ít sâu sắc hơn.

# Chương 7

## Đánh giá hệ thống

### 7.1 Đánh giá tính năng và hiệu quả của hệ thống

Hệ thống hỗ trợ học tập bằng cách sử dụng LLMs (Language Models) để tự động tạo ra các bài học đề xuất cho sinh viên dựa trên các bài học có sẵn từ giảng viên. Các tính năng đã triển khai, như trang dashboard, course list, course detail, và các trang liên quan đến việc đề xuất lessons đã bước đầu cho phép sinh viên tương tác hiệu quả với hệ thống học tập.

Trong quá trình phát triển, nhóm đã hoàn thành việc xây dựng giao diện và chức năng cơ bản, bao gồm:

- **Dashboard:** Cung cấp một cái nhìn tổng quan về các hoạt động học tập gần đây, tình trạng tiến độ học tập của sinh viên và các khóa học tham gia.
- **Course List và Course Detail:** Cho phép sinh viên theo dõi các khóa học, xem chi tiết các bài học, và tiếp cận với các tài liệu học tập.
- **Đề xuất Lessons:** Dựa trên mục tiêu học tập mà sinh viên đặt ra, hệ thống sẽ tự động tìm kiếm các bài học phù hợp từ các khóa học đã đăng tải, giúp sinh viên tiếp cận đúng bài học cần thiết.

Các mô-đun học tập (quiz, code exercises, reading material) đã được phát triển cho phép sinh viên tham gia vào việc học theo dạng quiz, code exercises, và tài liệu đọc. Hệ thống cũng đã ghi lại quá trình học tập của sinh viên, tạo điều kiện để tiến hành báo cáo và đánh giá kết quả học tập.



## 7.2 Đánh giá các hệ thống liên quan

Tiêu chí	Hệ thống Adaptive LearningPath [5]	Hệ thống FOKE [6]	CS50.ai[8]	APAS[9]	LeetCode
Cá nhân hóa	Đề xuất lộ trình dựa trên năng lực học sinh	Dựa trên hồ sơ người học và kỹ thuật prompt engineering	Gợi ý mã nguồn, trợ lý học tập AI	Phản hồi tức thời, gợi ý cải thiện mã	Tạo bài tập nhưng ít cá nhân hóa
Tính tương thích với người dùng	Thông qua đồ thị tri thức và tài nguyên phù hợp	Tùy chỉnh theo phong cách học tập và hành vi	Trợ lý học tập dựa trên GPT-4, trả lời câu hỏi học thuật	Phản hồi chi tiết về lỗi và cải tiến mã	Tạo bài tập dựa trên ngữ cảnh nhưng ít tùy chỉnh
Khả năng tương tác	Chỉ có gợi ý tài liệu học tập	Tương tác qua gợi ý cấu trúc và phản hồi liên tục	Chatbot tương tác trực tiếp với sinh viên	AI-Tutor cung cấp phản hồi và gợi ý cải thiện mã	Tương tác chủ yếu thông qua bài tập và câu hỏi
Độ chính xác của đề xuất	Tùy chỉnh theo dữ liệu thực tế	Sử dụng graph embedding và gợi ý cấu trúc	AI đưa ra phản hồi chính xác về mã nguồn	Phản hồi tức thời về mã nguồn và gợi ý cải thiện	Tạo bài tập nhưng không có giải thích chi tiết
Sự minh bạch và giải thích	Đưa ra lý do thay đổi lộ trình học	Cung cấp giải thích rõ ràng về lộ trình học	Giải thích chi tiết về mã nguồn và phong cách lập trình	Cung cấp lý do cài đặt mã, không tiết lộ đáp án	Ít giải thích về cách tạo bài tập hoặc câu hỏi
Tính mở rộng	Có thể mở rộng với nhiều tài nguyên học tập	Có thể mở rộng với nhiều đối tượng người học và lĩnh vực	Có thể áp dụng cho các khóa học online lớn	Có thể mở rộng cho nhiều khóa học và bài tập	Hệ thống có sẵn các bài tập cho nhiều cấp độ
Ứng dụng vào thực tế	Có thể áp dụng cho lớp học truyền thống và trực tuyến	Ứng dụng vào nhiều môi trường học tập khác nhau	Ứng dụng trong các khóa học lớn như CS50	Hệ thống đánh giá tự động có thể áp dụng cho lớp học và MOOCs	Ứng dụng cho các bài tập lập trình cá nhân và nhóm
Tiết kiệm thời gian cho giảng viên	Giảm tải công việc của giảng viên	Giảm tải cho giảng viên trong việc thiết kế lộ trình học	Giảm tải cho giảng viên trong việc giải thích mã	Giảm tải cho giảng viên trong việc kiểm tra và đánh giá bài tập	Giảm tải cho giảng viên trong việc tạo bài tập
Hiệu quả học tập	Cải thiện khả năng học của học sinh qua lộ trình cá nhân hóa	Cải thiện hiệu quả học tập nhờ cá nhân hóa sâu	Tăng cường khả năng tự học của sinh viên	Giúp sinh viên phát triển tư duy thuật toán	Hiệu quả học tập phụ thuộc vào việc tự học của người dùng

Table 7.1: Đánh giá các hệ thống liên quan

- Các tiêu chí đánh giá hệ thống

1. Cá nhân hóa: Đánh giá mức độ cá nhân hóa của lộ trình học tập, phản hồi và tài nguyên học tập.
2. Tính tương thích với người dùng: Đánh giá khả năng tương thích của hệ thống với các đặc điểm người dùng như phong cách học tập, mục tiêu học tập, và năng lực hiện tại.
3. Khả năng tương tác: Đánh giá mức độ và tính hiệu quả của sự tương tác giữa người học và hệ thống (ví dụ: chatbot, phản hồi tức thời).
4. Độ chính xác của đề xuất: Đánh giá khả năng của hệ thống trong việc tạo ra các đề xuất chính xác, phù hợp với người học.
5. Sự minh bạch và giải thích: Đánh giá mức độ giải thích rõ ràng của hệ thống về các đề xuất và lý do chọn các tài liệu học tập.
6. Tính mở rộng: Đánh giá khả năng mở rộng của hệ thống khi có nhiều người học và tài nguyên giáo dục hơn.



7. Ứng dụng vào thực tế: Đánh giá khả năng áp dụng hệ thống vào môi trường học tập thực tế, như lớp học truyền thống, lớp học trực tuyến, hoặc hệ thống MOOCs.
8. Tiết kiệm thời gian cho giảng viên: Đánh giá mức độ mà hệ thống giúp giảng viên giảm tải công việc như đánh giá bài tập, phản hồi học sinh, v.v.
9. Hiệu quả học tập: Đánh giá mức độ cải thiện kết quả học tập của sinh viên khi sử dụng hệ thống.

### 7.3 Các vấn đề và khó khăn trong quá trình phát triển

Mặc dù hệ thống đã đạt được nhiều tiến triển, nhưng quá trình phát triển không thiếu thử thách. Một số vấn đề chính bao gồm:

- **Khó khăn trong việc đồng bộ công nghệ:** Một số thành viên trong nhóm chưa thành thạo Docker, điều này gây khó khăn trong việc đồng bộ mã nguồn và triển khai hệ thống trên môi trường phát triển chung. Việc thiếu sự đồng thuận trong cách sử dụng các công cụ công nghệ cũng làm chậm tiến độ và dễ dẫn đến sự không nhất quán trong cấu hình hệ thống.
- **Vấn đề về thời gian của các thành viên:** Một số thành viên trong nhóm có lịch trình không đồng đều, điều này dẫn đến việc không thể hoàn thành các phần việc đúng hạn. Mặc dù đội ngũ đã cố gắng chia sẻ công việc hợp lý, nhưng sự thiếu hụt nhân lực vào những giai đoạn quan trọng khiến công việc bị trễ.
- **Chưa hoàn thiện phần chức năng Authentication:** Hiện tại, chức năng authentication của hệ thống chưa được hiện thực, khiến cho việc quản lý người dùng và bảo mật hệ thống chưa được đảm bảo. Điều này sẽ là một yếu tố quan trọng trong giai đoạn tiếp theo của hệ thống.
- **Chưa có kinh nghiệm kiểm thử chuyên nghiệp:** Mặc dù chúng tôi đã thực hiện một số kiểm thử API cơ bản, nhưng đội ngũ thiếu kinh nghiệm kiểm thử chuyên nghiệp và đầy đủ. Các khía cạnh khác như kiểm thử hiệu suất, giao diện người dùng (UI), và các thành phần của hệ thống vẫn chưa được kiểm tra một cách toàn diện, điều này có thể ảnh hưởng đến độ ổn định và hiệu quả của hệ thống trong các tình huống thực tế.
- **Tồn tại những testcase APIs failed trong quá trình phát triển:** Một số testcase APIs vẫn chưa hoàn thiện, điều này ảnh hưởng đến việc kiểm thử và đảm bảo chất lượng của hệ thống.
- **Chưa hoàn thiện tính năng tích hợp AI vào backend:** Mặc dù đã có sự nỗ lực trong việc tích hợp LLMs vào hệ thống, nhưng việc xây dựng mô-đun học tập cụ thể và sinh ra



các đề xuất bài học vẫn chưa được hoàn thiện. Điều này ảnh hưởng đến khả năng cung cấp các bài học chất lượng và phù hợp với nhu cầu học tập của sinh viên.

Ngoài ra, việc tích hợp LLMs để sinh ra các đề xuất bài học và xây dựng mô-đun học tập cụ thể cũng gặp một số thách thức về hiệu suất và độ chính xác của các đề xuất. Cần cải thiện và tối ưu thuật toán để nâng cao chất lượng các bài học sinh ra.

## **7.4 Kết chương**

Tổng thể, hệ thống học tập đã đạt được những kết quả bước đầu đáng ghi nhận, với nhiều tính năng hữu ích cho sinh viên. Tuy nhiên, còn nhiều vấn đề cần giải quyết, đặc biệt là liên quan đến việc đồng bộ công nghệ, hoàn thiện chức năng authentication, và cải thiện các đề xuất học tập thông qua LLMs. Những vấn đề này sẽ được khắc phục trong các giai đoạn phát triển tiếp theo.

# Chương 8

## Tổng kết

### 8.1 Kết quả đạt được

Sau một giai đoạn phát triển, hệ thống hỗ trợ học tập đã hoàn thành các tính năng cơ bản phục vụ cho việc học của sinh viên. Các phần quan trọng như trang dashboard, course list, course detail, và đề xuất bài học đã được triển khai đầy đủ, tạo ra một môi trường học tập chủ động và linh hoạt. Sinh viên có thể dễ dàng truy cập các khóa học, theo dõi tiến độ học tập, và nhận được các đề xuất học tập chính xác dựa trên mục tiêu cá nhân.

Các mô-đun học tập như quiz, code exercises, và reading material giúp sinh viên không chỉ học lý thuyết mà còn thực hành và tự kiểm tra khả năng của mình. Hệ thống cũng ghi lại các thống kê về thời gian học và tiến độ học tập, điều này giúp cung cấp các báo cáo chi tiết và đánh giá hiệu quả học tập của sinh viên.

### 8.2 Hướng phát triển

Mặc dù hệ thống hiện tại đã có những tính năng cơ bản, nhưng vẫn còn nhiều vấn đề cần phải giải quyết trong các giai đoạn tiếp theo. Cụ thể:

- Hoàn thiện phần authentication:** Đây là một yếu tố quan trọng để đảm bảo hệ thống có thể quản lý người dùng một cách an toàn và hiệu quả. Việc hoàn thiện chức năng này sẽ giúp phân quyền rõ ràng cho sinh viên, giảng viên, và quản trị viên (Admin).
- Tối ưu hóa công nghệ và quy trình làm việc:** Nhóm cần thống nhất việc sử dụng công cụ phát triển chung như Docker và các công nghệ hỗ trợ khác, đồng thời cải thiện quy trình làm việc để đảm bảo tiến độ phát triển không bị ảnh hưởng bởi vấn đề thời gian của các thành viên trong nhóm.
- Cải thiện tính chính xác của LLMs:** Hệ thống hiện tại cần được cải thiện về mặt độ



chính xác trong việc sinh ra các bài học và đề xuất phù hợp. Việc tối ưu hóa thuật toán sẽ giúp hệ thống gợi ý đúng bài học hơn, đồng thời giảm thiểu việc sinh ra các bài học không liên quan.

- **Mở rộng các tính năng học tập:** Cần bổ sung thêm các tính năng giảng viên và Admin để tạo một môi trường học tập sinh động và hiệu quả hơn.

Nhìn chung, hệ thống đang trên đà phát triển và sẽ còn nhiều công việc cần phải hoàn thiện trong giai đoạn tiếp theo để đáp ứng được đầy đủ các nhu cầu của người dùng.

## **8.3 Kế hoạch thực hiện Đồ án Tốt nghiệp**

### **Tuần 1**

**Tên nhiệm vụ:** Authentication: Backend APIs và UI

**Người đảm nhiệm:** Phạm Thi

**Chi tiết nhiệm vụ:**

- Phát triển API cho chức năng đăng nhập và đăng ký.
- Cài đặt hệ thống bảo mật cho API, bao gồm mã hóa mật khẩu và xác thực người dùng.
- Thiết kế giao diện UI cho trang đăng nhập và đăng ký.
- Kiểm tra tính năng đăng nhập và đăng ký với các trường hợp đặc biệt như mật khẩu sai, tài khoản không tồn tại.

**Tên nhiệm vụ:** Landing Page: Thiết kế UI cơ bản

**Người đảm nhiệm:** Tuấn Anh

**Chi tiết nhiệm vụ:**

- Thiết kế giao diện trang giới thiệu của hệ thống.
- Tích hợp thông tin cơ bản về hệ thống, cách thức hoạt động và các tính năng chính.
- Kiểm tra tính tương thích và khả năng tương tác của trang giới thiệu.

**Tên nhiệm vụ:** Quản lý khóa học (Giảng viên): API CRUD

**Người đảm nhiệm:** Phương Nam

**Chi tiết nhiệm vụ:**

- Phát triển các API cho phép giảng viên thêm, sửa và xóa các khóa học.



- Đảm bảo các API này có thể xử lý các tình huống dữ liệu như khóa học đã có, khóa học không hợp lệ.
- Phát triển tính năng yêu cầu giảng viên upload tài liệu cho mỗi bài học (lesson) trong khóa học. Mỗi lesson cần có ít nhất một tài liệu PDF (slides của chapter) được upload. Giảng viên có thể upload các loại tài liệu khác, nhưng sẽ có các giới hạn về định dạng file (ví dụ: chỉ cho phép các định dạng như .pdf, .docx, .pptx, v.v.).
- Đảm bảo rằng giảng viên phải nhập tên và thứ tự của từng lesson (chapter). Trường mô tả (description) là tùy chọn, không bắt buộc.
- Kiểm tra API với các bộ dữ liệu mẫu, bao gồm các trường hợp không hợp lệ như tải lên tài liệu không đúng định dạng, thiếu tên hoặc thứ tự bài học.
- Cung cấp các thông báo lỗi rõ ràng nếu tài liệu không hợp lệ hoặc thiếu thông tin cần thiết cho bài học.
- Kiểm tra API với các tình huống thực tế để đảm bảo tính chính xác, hiệu suất và khả năng mở rộng của hệ thống khi có nhiều bài học và tài liệu được thêm vào.

## Tuần 2-3

**Tên nhiệm vụ:** Quản lý học sinh (Giảng viên): API + UI

**Người đảm nhiệm:** Phương Nam

**Chi tiết nhiệm vụ:**

- Phát triển API cho phép giảng viên xem tiến độ học tập của học sinh.
- Thiết kế giao diện UI cho phép giảng viên dễ dàng xem và tương tác với tiến độ của từng học sinh.
- Thêm các tính năng lọc, tìm kiếm và phân loại học sinh theo tiến độ học tập.
- Kiểm tra tính năng giao diện và API.

**Tên nhiệm vụ:** Progress Tracking: Backend APIs

**Người đảm nhiệm:** Phạm Thi

**Chi tiết nhiệm vụ:**

- Phát triển API để xử lý và lưu trữ dữ liệu học tập của sinh viên.
- Đảm bảo tính chính xác trong việc theo dõi tiến độ học tập của sinh viên qua các bài học và khóa học.



- Tích hợp các API với các hệ thống khác để đảm bảo dữ liệu được đồng bộ và cập nhật đúng lúc.

**Tên nhiệm vụ:** Quản lý người dùng (Admin): API CRUD

**Người đảm nhiệm:** Tuấn Anh

**Chi tiết nhiệm vụ:**

- Phát triển API cho phép quản lý tài khoản của giảng viên và sinh viên.
- Thiết lập quyền truy cập cho admin để có thể xem, thêm, sửa và xóa tài khoản người dùng.
- Đảm bảo các API bảo mật thông tin người dùng và thực hiện các kiểm tra bảo mật.

## **Tuần 4-6**

**Tên nhiệm vụ:** Tích hợp LLM: Quiz Generation, Document Generation (Backend APIs)

**Người đảm nhiệm:** Phạm Thi

**Chi tiết nhiệm vụ:**

- Phát triển API sử dụng LLM (Large Language Model) để tạo các câu hỏi trắc nghiệm tự động dựa trên bài giảng.
- Tạo tài liệu tự động từ nội dung khóa học, bao gồm phần tóm tắt và các câu hỏi ôn tập.
- Kiểm tra độ chính xác của các câu hỏi được tạo ra và tính hiệu quả của tài liệu tự động.

**Tên nhiệm vụ:** Tích hợp LLM: Code Hint Module, Debugging Assistant, Explain Code Module

**Người đảm nhiệm:** Tuấn Anh

**Chi tiết nhiệm vụ:**

- Phát triển API cung cấp gợi ý sửa lỗi và giải thích code tự động khi sinh viên gặp vấn đề.
- Tích hợp module giải thích các lỗi và cung cấp hướng dẫn sửa chữa trong quá trình lập trình.
- Thực hiện kiểm tra và tối ưu hóa các module này để đảm bảo tính chính xác và hiệu quả.

**Tên nhiệm vụ:** Xử lý câu trả lời: LLM Evaluation Module

**Người đảm nhiệm:** Phương Nam

**Chi tiết nhiệm vụ:**

- Phát triển API chấm điểm tự động cho câu trả lời của sinh viên.



- Sử dụng LLM để phân tích và đưa ra nhận xét về câu trả lời của sinh viên.
- Đảm bảo API có thể xử lý các loại câu hỏi khác nhau và đưa ra đánh giá chính xác.

## Tuần 7-8

**Tên nhiệm vụ:** Progress Tracking: UI + Biểu đồ

**Người đảm nhiệm:** Phạm Thị

**Chi tiết nhiệm vụ:**

- Thiết kế giao diện UI hiển thị tiến độ học tập của sinh viên thông qua biểu đồ và đồ thị.
- Đảm bảo rằng giao diện có thể cập nhật và phản hồi nhanh chóng với dữ liệu mới.
- Kiểm tra tính tương thích của giao diện với các thiết bị và nền tảng khác nhau.

**Tên nhiệm vụ:** Gửi Feedback (Sinh viên): Backend APIs

**Người đảm nhiệm:** Phương Nam

**Chi tiết nhiệm vụ:**

- Phát triển API cho phép sinh viên gửi phản hồi về bài học và hệ thống.
- Đảm bảo hệ thống lưu trữ và xử lý phản hồi của sinh viên một cách chính xác và hiệu quả.
- Thực hiện các kiểm tra và tối ưu hóa để đảm bảo tính ổn định và bảo mật cho API.

**Tên nhiệm vụ:** Hiển thị Recent Activities: Backend APIs

**Người đảm nhiệm:** Tuấn Anh

**Chi tiết nhiệm vụ:**

- Phát triển API lưu trữ và hiển thị các hoạt động gần đây của sinh viên.
- Đảm bảo API này có thể xử lý và truy vấn dữ liệu một cách nhanh chóng và chính xác.
- Kiểm tra tính năng API trong các tình huống thực tế và đảm bảo dữ liệu không bị mất hoặc sai sót.

## Tuần 9-10

**Tên nhiệm vụ:** Gửi Feedback: UI tích hợp

**Người đảm nhiệm:** Phương Nam

**Chi tiết nhiệm vụ:**



- Thiết kế giao diện người dùng cho tính năng gửi phản hồi của sinh viên.
- Cung cấp các tùy chọn cho sinh viên để chọn loại phản hồi và nhập ý kiến.
- Tối ưu giao diện để đảm bảo tính thân thiện và dễ sử dụng trên nhiều nền tảng khác nhau.
- Kiểm tra các chức năng của giao diện để đảm bảo dữ liệu phản hồi được gửi đúng và đầy đủ.

**Tên nhiệm vụ:** Hiển thị Recent Activities: UI tích hợp

**Người đảm nhiệm:** Tuấn Anh

**Chi tiết nhiệm vụ:**

- Thiết kế giao diện UI để hiển thị các hoạt động gần đây của sinh viên, bao gồm các bài học, bài tập, và phản hồi.
- Cung cấp khả năng lọc và tìm kiếm các hoạt động theo thời gian hoặc loại hoạt động.
- Đảm bảo giao diện có thể cập nhật động khi có dữ liệu mới từ các API backend.
- Kiểm tra giao diện với các tình huống thực tế để đảm bảo hiển thị chính xác và dễ sử dụng.

**Tên nhiệm vụ:** Quản lý hệ thống (Admin): Logs & Lỗi

**Người đảm nhiệm:** Tuấn Anh

**Chi tiết nhiệm vụ:**

- Phát triển API và giao diện người dùng cho phép admin theo dõi các logs và lỗi hệ thống.
- Xây dựng chức năng cảnh báo khi phát hiện lỗi nghiêm trọng hoặc sự cố trong hệ thống.
- Đảm bảo tính năng có thể phân loại các lỗi và sự cố theo mức độ ưu tiên và loại hình.
- Kiểm tra hiệu quả và độ chính xác của tính năng giám sát lỗi trong các tình huống thực tế.

## **Tuần 11-12**

**Tên nhiệm vụ:** Xử lý Feedback (Admin): Backend APIs

**Người đảm nhiệm:** Phương Nam

**Chi tiết nhiệm vụ:**

- Phát triển API cho phép admin xử lý phản hồi từ sinh viên.



- Tích hợp hệ thống để admin có thể xem, phân loại và trả lời phản hồi của sinh viên.
- Đảm bảo API có khả năng phân tích các phản hồi tự động và cung cấp các giải pháp hoặc hướng dẫn.
- Kiểm tra tính hiệu quả của hệ thống trong việc quản lý và xử lý phản hồi.

**Tên nhiệm vụ:** Quản lý bài tập (Giảng viên): API CRUD

**Người đảm nhiệm:** Phạm Thi

**Chi tiết nhiệm vụ:**

- Phát triển các API cho phép giảng viên quản lý bài tập trong các khóa học.
- Đảm bảo giảng viên có thể tạo, chỉnh sửa, xóa và xem các bài tập.
- Xây dựng API để giảng viên có thể theo dõi tình trạng hoàn thành bài tập của sinh viên.
- Kiểm tra tính ổn định và chính xác của API trong các tình huống sử dụng thực tế.

**Tên nhiệm vụ:** Testing Feedback Module

**Người đảm nhiệm:** Phương Nam

**Chi tiết nhiệm vụ:**

- Kiểm tra và debug toàn bộ module Feedback để đảm bảo tính năng hoạt động đúng.
- Chạy thử nghiệm với nhiều tình huống khác nhau để kiểm tra độ chính xác và tính ổn định.
- Xử lý các lỗi phát sinh và tối ưu hóa các chức năng liên quan đến phản hồi.
- Kiểm tra khả năng xử lý khối lượng lớn dữ liệu phản hồi từ nhiều người dùng cùng lúc.

## Tuần 13

**Tên nhiệm vụ:** Testing & Debugging: Toàn bộ hệ thống

**Người đảm nhiệm:** Cả nhóm

**Chi tiết nhiệm vụ:**

- Kiểm tra toàn bộ hệ thống để phát hiện và sửa lỗi.
- Thực hiện các bài test tự động và thủ công trên cả backend và frontend.
- Đảm bảo hệ thống vận hành mượt mà với tất cả các tính năng được tích hợp.
- Xử lý các vấn đề hiệu suất và bảo mật nếu có.
- Kiểm tra khả năng mở rộng của hệ thống khi có nhiều người dùng truy cập đồng thời.



## Tuần 14

**Tên nhiệm vụ:** Triển khai Backend: Deploy APIs

**Người đảm nhiệm:** Phạm Thị

**Chi tiết nhiệm vụ:**

- Triển khai các API backend lên môi trường test hoặc production.
- Cấu hình các server và cơ sở hạ tầng để chạy API một cách ổn định.
- Kiểm tra kết nối và hiệu suất của các API sau khi triển khai.
- Đảm bảo hệ thống có thể xử lý lưu lượng truy cập thực tế từ người dùng.

**Tên nhiệm vụ:** Triển khai Frontend: Deploy UI

**Người đảm nhiệm:** Trần Anh

**Chi tiết nhiệm vụ:**

- Triển khai giao diện người dùng lên môi trường test hoặc production.
- Đảm bảo giao diện hoạt động mượt mà và không gặp lỗi trên các trình duyệt và thiết bị khác nhau.

## Tuần 15

**Tên nhiệm vụ:** Báo cáo & Demo: Chuẩn bị thuyết trình

**Người đảm nhiệm:** Cả nhóm

**Chi tiết nhiệm vụ:**

- Chuẩn bị báo cáo đầy đủ về quá trình phát triển dự án, các tính năng đã hoàn thành và kết quả đạt được.
- Tạo slide thuyết trình để trình bày các điểm nổi bật của hệ thống và quy trình làm việc.
- Thực hiện demo hệ thống cho các đối tượng liên quan, bao gồm giảng viên và sinh viên.
- Sửa lỗi và cải thiện các tính năng nếu cần thiết dựa trên phản hồi từ người dùng trong buổi demo.

# Bibliography

- [1] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. <https://arxiv.org/abs/2005.14165>
- [2] IBM. (n.d.). What is a knowledge graph? [Accessed: 2024-09-22]. <https://www.ibm.com/topics/knowledge-graph>
- [3] Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., & Sohl-Dickstein, J. (2015). Deep knowledge tracing. *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*, 505–513.
- [4] Abdelrahman, G., Wang, Q., & Nunes, B. (2023). Knowledge tracing: A survey. *ACM Comput. Surv.*, 55(11). <https://doi.org/10.1145/3569576>
- [5] Raj, N. S., & Renumol, V. G. (2022). An improved adaptive learning path recommendation model driven by real-time learning analytics. *Journal of Computers in Education*, 11, 121–148. <https://doi.org/10.1007/s40692-022-00250-y>
- [6] Hu, S., & Wang, X. (2024). Foke: A personalized and explainable education framework integrating foundation models, knowledge graphs, and prompt engineering. <https://arxiv.org/abs/2405.03734>
- [7] Ng, C., & Fung, Y. (2024). Educational personalized learning path planning with large language models. <https://arxiv.org/abs/2407.11773>
- [8] Liu, R., Zenke, C., Liu, C., Holmes, A., Thornton, P., & Malan, D. J. (2024). Teaching cs50 with ai: Leveraging generative artificial intelligence in computer science education. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V*, 1, 750–756. <https://doi.org/10.1145/3626252.3630938>
- [9] Frankford, E., Sauerwein, C., Bassner, P., Krusche, S., & Breu, R. (2024). Ai-tutoring in software engineering education. *Proceedings of the 46th International Conference on Software Engineering: Software Engineering Education and Training*, 309–319. <https://doi.org/10.1145/3639474.3640061>
- [10] Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A systematic survey of prompt engineering in large language models: Techniques and applications.



- [11] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Chi, E. H., Le, Q., & Zhou, D. (2022). Chain of thought prompting elicits reasoning in large language models. *CoRR*, *abs/2201.11903*. <https://arxiv.org/abs/2201.11903>
- [12] Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2024). Retrieval-augmented generation for large language models: A survey. <https://arxiv.org/abs/2312.10997>
- [13] Lin, D. (2024). Revolutionizing retrieval-augmented generation with enhanced pdf structure recognition. <https://arxiv.org/abs/2401.12599>
- [14] Bulathwela, S., Perez-Ortiz, M., Ranawaka, E. S. V., Siriwardana, R. I. P. B. B., Ganepola, G. A. K. Y., Ravikkumar, T., Pussegoda, S., Novak, E., Yilmaz, E., & Shawe-Taylor, J. (2021). Cross modal, cross cultural, cross lingual, cross domain, and cross site global oer network.
- [15] Li, H., Yu, J., Ouyang, Y., Liu, Z., Rong, W., Li, J., & Xiong, Z. (2024). Explainable few-shot knowledge tracing. *arXiv preprint arXiv:2405.14391*.