

TRƯỜNG ĐẠI HỌC GIAO THÔNG VẬN TẢI
PHÂN HIỆU TẠI TP. HỒ CHÍ MINH
BỘ MÔN CÔNG NGHỆ THÔNG TIN



BÁO CÁO ĐỒ ÁN TỐT NGHIỆP

ĐỀ TÀI: NGHIÊN CỨU THUẬT TOÁN KMEANS TRONG
GOM CỤM DỮ LIỆU, ỨNG DỤNG PHÁT TRIỂN LMS TÍCH
HỢP KIỂM TRA ĐẠO VĂN

Sinh viên thực hiện: **PHẠM THÀNH HẬU**

Lớp : **CÔNG NGHỆ THÔNG TIN K59**

Khoá : **K59**

Tp. Hồ Chí Minh, năm 2023

TRƯỜNG ĐẠI HỌC GIAO THÔNG VẬN TẢI
PHÂN HIỆU TẠI TP. HỒ CHÍ MINH
BỘ MÔN CÔNG NGHỆ THÔNG TIN



BÁO CÁO ĐỒ ÁN TỐT NGHIỆP

ĐỀ TÀI: NGHIÊN CỨU THUẬT TOÁN KMEANS TRONG
GOM CỤM DỮ LIỆU, ỨNG DỤNG PHÁT TRIỂN LMS TÍCH
HỢP KIỂM TRA ĐẠO VĂN

Sinh viên thực hiện: **PHẠM THÀNH HẬU**

Lớp : **CÔNG NGHỆ THÔNG TIN K59**

Khoá : **K59**

Tp. Hồ Chí Minh, năm 2023

BỘ MÔN: CÔNG NGHỆ THÔNG TIN

THIẾT KẾ TỔNG QUAN ĐỀ TÀI

-----***-----

Mã sinh viên : 5951071024

Họ tên sinh viên : Phạm Thành Hậu

Khóa: 59 Lớp: CQ.59.CNTT

1. Tên đề tài

Nghiên cứu thuật toán kmeans trong gom cụm dữ liệu, ứng dụng phát triển LMS tích hợp kiểm tra đạo văn.

2. Mục đích, yêu cầu

a. Mục đích:

Áp dụng thuật toán kmeans gom cụm dữ liệu để tối ưu tốc độ kiểm tra đạo văn và xây dựng ứng dụng LMS

b. Yêu cầu:

➤ Yêu cầu công nghệ

- Sử dụng ngôn ngữ lập trình : Javascript, HTML, Ajax, JQuery, JavaScript, Bootstrap, C#, golang.
- Sử dụng công cụ Visual Studio 2019, SQL Server và .Net Framework, echo Framework.
- Công nghệ sử dụng ASP.NET MVC.

➤ Yêu cầu chức năng

- Tạo lớp, tham gia lớp.
- Đăng thông báo ,bài tập.
- Nộp và chấm điểm bài tập.
- Xem điểm và mức độ hoàn thành bài tập.
- Nhắn tin.
- Kiểm tra đạo văn.
- Gom cụm dữ liệu

➤ Yêu cầu phi chức năng

- Tốc độ: Tối ưu về tốc độ, làm việc với dữ liệu vừa phải, tốc độ hiển thị nhanh.
- Giao diện: Thân thiện với người dùng và dễ dàng thao tác.

3. Nội dung và phạm vi đề tài

a. Nội dung:

- Tổng quan và lý do chọn đề tài nghiên cứu thuật toán kmeans trong gom cụm dữ liệu, ứng dụng phát triển LMS tích hợp kiểm tra đạo văn..
 - Mục tiêu đề tài
 - Cơ sở lý thuyết
 - Phân tích thiết kế hệ thống.
- b. Phạm vi:**
- Nghiên cứu sử dụng công cụ Visual Studio ,ngôn ngữ golang và ngôn ngữ C#.
 - Nghiên cứu xây dựng website LMS có nhấn tin tìm tài liệu,sử lý theo thời gian thực và kiểm tra đạo văn.
- 4. Công nghệ, công cụ và ngôn ngữ lập trình**
- Công cụ lập trình: Visual Studio 2019, SQL Server và .Net Framework, echo Framework.
 - Công nghệ sử dụng: ASP.NET MVC.
 - Ngôn ngữ lập trình: Javascript,HTML,Ajax,Jquery,JavaScript, Bootstrap, C#, golang.
- 5. Các kết quả chính dự kiến sẽ đạt được và ứng dụng**
- Bài báo cáo đề tài (bản Word và File Power Point)
 - Xây dựng được website LMS có kiểm tra đạo văn và gom cụm dữ liệu cho việc tối ưu hóa tốc độ kiểm tra đạo văn .
- 6. Giảng viên và cán bộ hướng dẫn**
- Họ tên: Trần Phong Nhã
- Đơn vị công tác: Bộ môn Công nghệ Thông tin - Trường Đại học Giao thông Vận tải Phân hiệu tại thành phố Hồ Chí Minh.
- Điện thoại:

Ngày ... tháng năm 2023

Trưởng BM Công nghệ Thông tin

Đã giao nhiệm vụ đồ án tốt nghiệp

Giảng viên hướng dẫn

ThS. Trần Phong Nhã

ThS. Trần Phong Nhã

LỜI CẢM ƠN

Để hoàn thành đề tài này trước hết em xin gửi đến quý thầy, cô Bộ môn Công nghệ thông tin – Phân hiệu Trường Đại học Giao thông Vận tải tại Thành phố Hồ Chí Minh lời cảm ơn chân thành vì đã truyền đạt cho em những kiến thức không chỉ từ sách vở, mà còn những kinh nghiệm quý giá từ cuộc sống trong khoảng thời gian học tập tại trường.

Tiếp đến, em xin trân trọng bày tỏ sự cảm ơn chân thành nhất đến thầy Trần Phong Nhã giảng viên bộ môn Công Nghệ Thông Tin. Thầy là người hướng dẫn chính. Các định hướng cùng những phương pháp và lời góp ý của Thầy là yếu tố quan trọng giúp chúng em hoàn thành đề tài này.

Vì hiểu biết cá nhân còn hạn chế, em cũng đã nỗ lực hết sức để hoàn thành đề tài một cách tốt nhất, nhưng chắc chắn vẫn sẽ có những thiếu sót không thể tránh khỏi. Em kính mong nhận được sự thông cảm và những ý kiến đóng góp chân thành từ quý thầy cô.

Sau cùng, em xin kính chúc thầy Trần Phong Nhã cũng như Quý Thầy Cô trong Bộ môn Công nghệ thông tin hạnh phúc và thành công hơn nữa trong công việc cũng như trong cuộc sống.

Em xin chân thành cảm ơn!

Tp. Hồ Chí Minh, ngày tháng năm 2023

Sinh viên thực hiện

Phạm Thành Hậu

NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Tp. Hồ Chí Minh, ngày tháng năm 2023

Giảng viên hướng dẫn

ThS. Trần Phong Nhã

MỤC LỤC

THIẾT KẾ TỔNG QUAN ĐỀ TÀI.....	i
MỤC LỤC	v
DANH MỤC HÌNH ẢNH.....	viii
DANH MỤC BẢNG BIỂU	x
MỞ ĐẦU	1
1. Lý do chọn đề tài	1
2. Mục tiêu của đề tài.....	1
3. Lựa chọn công cụ cài đặt.....	2
4. Phạm vi đề tài	2
CHƯƠNG 1: CƠ SỞ LÝ THUYẾT.....	3
1.1. Công nghệ sử dụng	3
1.1.1. ASP.NET MVC.....	3
1.1.2. Bootstrap.....	4
1.1.3. Signalr	5
1.1.4. Spire.office	6
1.1.5. Cronjob.....	6
1.1.6. Echo.....	7
1.2. Ngôn ngữ lập trình.....	8
1.2.1. C#	8
1.2.2. Javascript.....	9
1.2.3. Giới thiệu về HTML.....	10
1.2.4. Giới thiệu về CSS.....	11
1.2.5. Golang	12
1.3. Cơ sở dữ liệu.....	12
1.3.1. Hệ quản trị cơ sở dữ liệu	12
1.3.2. Cơ sở dữ liệu	13
1.4. Thuật toán gom cụm dữ liệu.....	14
1.4.1. K-means	14
1.4.2. Agglomerative Hierarchical Clustering	15
1.4.2. Density-Based Spatial Clustering of Applications with Noise	16
1.4.2. Mean-Shift.....	17
1.5. Thuật toán áp dụng	19

1.5.1. Thuật toán Kmeans.....	19
1.5.2. Thuật toán kiểm tra đạo văn.....	21
CHƯƠNG 2: PHÂN TÍCH HỆ THỐNG	24
2.1. Mô tả bài toán.....	24
2.2. Sơ đồ phân cấp chức năng BFD	26
2.3. Bảng phân tích xác định tiến trình, tác nhân và hồ sơ.....	26
2.3.1. Xác định tiến trình.....	26
2.3.2. Tác nhân và hồ sơ.....	28
2.4. Biểu đồ luồng dữ liệu (mức ngưỡng cảnh, mức đỉnh, mức dưới đỉnh)	28
2.4.1.DFD mức ngưỡng cảnh.....	28
2.4.2.DFD mức đỉnh	29
2.4.3.DFD mức dưới đỉnh	30
2.5. Biểu đồ UseCase.....	33
2.5.1. Tổng quan hệ thống.....	33
2.5.2. Các tác nhân	34
2.5.3. Sơ đồ UseCase tổng quát.....	35
2.6. Biểu đồ hoạt động.....	35
2.6.1. Chức năng tạo lớp học.....	35
2.6.2. Chức năng tham gia lớp học.....	36
2.6.3. Chức năng tạo bài tập tự luận.....	36
2.6.4. Chức năng tạo bài tập trắc nghiệm.....	37
2.6.5. Chức năng tạo bài tập trắc nghiệm.....	37
2.6.6. Chức năng tìm kiếm tài liệu	38
2.6.7. Chức năng nhắn tin	38
2.6.8. Chức năng làm bài tập trắc nghiệm.....	39
2.6.9. Chức năng làm bài tập tự luận.....	39
2.7 . Sơ đồ tuần tự.....	40
2.7.1. Sơ đồ tuần tự tìm kiếm tài liệu	40
2.7.2. Sơ đồ tuần tự kiểm tra đạo văn.....	40
2.7.3. Sơ đồ tuần tự tạo lớp học và tham gia lớp học.....	41
2.7.4. Sơ đồ tuần tự làm bài tập trắc nghiệm.....	42
2.7.5. Sơ đồ tuần tự chấm bài tập tự luận.....	42
CHƯƠNG 3: THIẾT KẾ HỆ THỐNG VÀ TRIỂN KHAI.....	43
3.1. Thiết kế hệ thống	43
3.1.1. Các thực thể và thuộc tính.....	43
3.1.2. Xây dựng mô hình thực thể liên kết.....	50

3.3.3. Mô hình thực thể liên kết sang mô hình quan hệ	52
3.2 Triển khai.....	57
3.2.1 Giao diện đăng nhập.....	57
3.2.2 Giao diện trang chủ	57
3.2.5 Giao diện thông báo	58
3.2.7 Giao diện bài tập	58
3.2.8 Giao diện thành viên lớp học	64
3.2.9 Giao diện thống kê điểm	65
3.2.10 Giao diện tài liệu	66
3.2.11 Giao diện nhắn tin	66
3.2.12 Giao diện thông tin người dùng	67
3.2.13 Giao diện quản lý của Admin.....	68
KẾT LUẬN VÀ KIẾN NGHỊ	72
1.Kết quả đạt được.....	72
2.Tồn tại.....	72
3.Hướng phát triển.....	72
TÀI LIỆU THAM KHẢO	73

DANH MỤC HÌNH ẢNH

Hình 2. 1 Sơ đồ phân cấp chức năng	26
Hình 2. 2 DFD mức ngưỡng cảnh.	28
Hình 2. 3 DFD mức đỉnh.	29
Hình 2. 4 DFD mức dưới đỉnh của quản lý tài khoản.....	30
Hình 2. 5 DFD mức dưới đỉnh của quản lý lớp học.....	30
Hình 2. 6 DFD mức dưới đỉnh của quản lý thông báo.....	31
Hình 2. 7 DFD mức dưới đỉnh của quản lý bài tập.....	31
Hình 2. 8 DFD mức dưới đỉnh của thống kê điểm	32
Hình 2. 9 DFD mức dưới đỉnh của nhắn tin.	32
Hình 2. 10 DFD mức dưới đỉnh của tài liệu.	33
Hình 2. 11 Sơ đồ Usecase tổng quát	35
Hình 2. 12 Biểu đồ hoạt động tạo lớp học.....	35
Hình 2. 13 Biểu đồ hoạt động tham gia lớp học	36
Hình 2. 14 Biểu đồ hoạt động tạo bài tập tự luận.....	36
Hình 2. 15 Biểu đồ hoạt động tạo bài tập trắc nghiệm.....	37
Hình 2. 16 Biểu đồ hoạt động kiểm tra đạo văn.....	37
Hình 2. 17 Biểu đồ hoạt động tìm kiếm tài liệu.....	38
Hình 2. 18 Biểu đồ hoạt động nhắn tin	38
Hình 2. 19 Biểu đồ hoạt động làm bài trắc nghiệm	39
Hình 2. 20 Biểu đồ hoạt động làm bài tự luận.....	39
Hình 2. 21 Sơ đồ tuần tự tìm kiếm tài liệu.....	40
Hình 2. 22 Sơ đồ tuần tự kiểm tra đạo văn.....	40
Hình 2. 23 Sơ đồ tuần tự tạo lớp và tham gia lớp học	41
Hình 2. 24 Sơ đồ tuần tự làm bài trắc nghiệm	42
Hình 2. 25 Sơ đồ tuần tự chấm bài tự luận	42
Hình 3. 1 Mô hình thực thể liên kết	51
Hình 3. 2 Giao diện đăng nhập	57
Hình 3. 3 Giao diện trang chủ.....	58
Hình 3. 4 Giao diện thông báo	58
Hình 3. 5 Giao diện bài tập cho giảng viên	59
Hình 3. 6 Giao diện điểm bài tập trắc nghiệm.....	59
Hình 3. 7 Giao diện điểm bài tập tự luận.....	60
Hình 3. 8 Giao diện kiểm tra đạo văn	60
Hình 3. 9 Giao diện câu hỏi trắc nghiệm	61
Hình 3. 10 Giao diện câu hỏi tự luận.....	61
Hình 3. 11 Giao diện bài tập trắc nghiệm.....	62

<i>Hình 3. 12 Giao diện bài trắc nghiệm.....</i>	<i>62</i>
<i>Hình 3. 13 Giao diện bài tập tự luận.....</i>	<i>63</i>
<i>Hình 3. 14 Giao diện tạo bài tập trắc nghiệm.....</i>	<i>63</i>
<i>Hình 3. 15 Giao diện bài tập tự luận.....</i>	<i>64</i>
<i>Hình 3. 16 Giao diện bài tập cho sinh viên.....</i>	<i>64</i>
<i>Hình 3. 17 Giao diện thành viên lớp học</i>	<i>65</i>
<i>Hình 3. 18 Giao diện thống kê điểm.....</i>	<i>65</i>
<i>Hình 3. 19 Giao diện tài liệu.....</i>	<i>66</i>
<i>Hình 3. 20 Giao diện xem tài liệu.....</i>	<i>66</i>
<i>Hình 3. 21 Giao diện nhắn tin.....</i>	<i>67</i>
<i>Hình 3. 22 Giao diện thông tin người dùng</i>	<i>67</i>
<i>Hình 3. 23 Giao diện thay đổi mật khẩu</i>	<i>68</i>
<i>Hình 3. 24 Giao diện trang chủ Admin</i>	<i>68</i>
<i>Hình 3. 25 Giao diện quản lý lớp học dành cho Admin.....</i>	<i>69</i>
<i>Hình 3. 26 Giao diện quản lý thành viên lớp học dành cho Admin</i>	<i>69</i>
<i>Hình 3. 27 Giao diện quản lý cụm dữ liệu</i>	<i>70</i>
<i>Hình 3. 28 Giao diện quản lý chi tiết cụm dữ liệu</i>	<i>70</i>
<i>Hình 3. 29 Giao diện quản lý tài khoản</i>	<i>71</i>

DANH MỤC BẢNG BIỂU

<i>Bảng 1. 1 Bảng danh sách tài liệu.....</i>	<i>20</i>
<i>Bảng 1. 2 Bảng danh sách từ.....</i>	<i>20</i>
<i>Bảng 1. 3 Bảng danh sách cụm từ.....</i>	<i>20</i>
<i>Bảng 1. 4 Bảng danh sách điểm trung tâm</i>	<i>20</i>
<i>Bảng 1. 5 Bảng danh sách cụm điểm</i>	<i>21</i>
<i>Bảng 1. 6 Bảng danh sách câu</i>	<i>22</i>
<i>Bảng 1. 7 Bảng phân trăm câu.....</i>	<i>22</i>
<i>Bảng 1. 8 Bảng danh sách cụm từ.....</i>	<i>22</i>
<i>Bảng 1. 9 Bảng danh sách phân trăm</i>	<i>23</i>
<i>Bảng 2. 1 Bảng xác định tiến trình.....</i>	<i>28</i>
<i>Bảng 2. 2 Các tác nhân</i>	<i>34</i>
<i>Bảng 3. 1 Thực thể tài khoản.....</i>	<i>43</i>
<i>Bảng 3. 2 Thực thể lớp học.....</i>	<i>44</i>
<i>Bảng 3. 3 Thực thể thông báo.....</i>	<i>44</i>
<i>Bảng 3. 4 Thực thể tin nhắn.</i>	<i>44</i>
<i>Bảng 3. 5 Thực thể file thông báo.</i>	<i>45</i>
<i>Bảng 3. 6 Thực thể bài tập.</i>	<i>45</i>
<i>Bảng 3. 7 Thực thể file bài tập tự luận.....</i>	<i>46</i>
<i>Bảng 3. 8 Thực thể thông tin bài tập tự luận.....</i>	<i>46</i>
<i>Bảng 3. 9 Thực thể đáp án.....</i>	<i>46</i>
<i>Bảng 3. 10 Thực thể câu hỏi.....</i>	<i>47</i>
<i>Bảng 3. 11 Thực thể bài tập trắc nghiệm.</i>	<i>47</i>
<i>Bảng 3. 12 Thực thể bài tập tự luận.</i>	<i>48</i>
<i>Bảng 3. 13 Thực thể bình luận bài tập tự luận.....</i>	<i>48</i>
<i>Bảng 3. 14 Thực thể bình luận bài tập trắc nghiệm.....</i>	<i>48</i>
<i>Bảng 3. 15 Thực thể bình luận thông báo.</i>	<i>49</i>
<i>Bảng 3. 16 Thực thể bình luận bài tập tự luận.....</i>	<i>49</i>
<i>Bảng 3. 17 Thực thể bình luận bài tập trắc nghiệm.....</i>	<i>49</i>
<i>Bảng 3. 18 Thực thể trả lời bình luận thông báo.</i>	<i>50</i>
<i>Bảng 3. 19 Thực thể kiểm tra đạo văn.....</i>	<i>50</i>
<i>Bảng 3. 20 Thực thể tài liệu.</i>	<i>50</i>

MỞ ĐẦU

1. Lý do chọn đề tài

Hiện nay, thời đại công nghệ số ngày càng phát triển nên hình thức giao và nộp bài tập online ngày càng phổ biến hơn. Công cụ được nhiều người sử dụng phổ biến đó là classroom do google phát hành. Tuy nhiên khi sử dụng ứng dụng thì gặp nhiều vấn đề như khi muốn hỏi bài thành viên trong lớp thì phải sử dụng ứng dụng khác và tìm tài liệu phải tìm ngoài những nguồn khác. Nhằm hỗ trợ việc nâng cao lượng thông tin bài học vững chắc hơn cũng như những thông tin muốn tìm hiểu thêm kiến thức.

Với những lý do đã nêu trên, em đã xây dựng một ứng dụng web với chức năng hỗ trợ trong vấn đề học tập.

2. Mục tiêu của đề tài

Hệ thống phải có giao diện phù hợp, dễ sử dụng, thông tin lưu trữ một cách tối ưu. Các chức năng phải sát với yêu cầu thực tế của hình thức bài tập trực tuyến. Hệ thống có khả năng hỗ trợ đa người dùng, độ bảo mật cao. Với những đặc điểm như vậy, hệ thống thực hiện các công việc như sau :

- Xây dựng thống có chức năng tạo lớp học ,tham gian lớp học.
- Đăng thông báo sử dụng công nghệ runtime khi một người dùng đăng thông báo thì những người khác khi đang ở trang thông báo sẽ hiện yêu cầu reload lại trang.
- Đăng bài tập sử dụng công nghệ runtime khi giảng viên đăng bài tập thì những thành viên trong lớp khi đang ở trang bài tập sẽ hiện yêu cầu reload lại trang.
- Giảng viên chấm điểm bài tập đối với bài tập tự luận. Còn bài tập trắc nghiệm thì khi sinh viên làm bài xong hệ thống sẽ tự chấm điểm và đưa ra kết quả
- Giảng viên có thể xem điểm trung bình của từng sinh viên hiện số lượng bài tập đã nộp chưa nộp hoặc nộp muộn. Sinh viên có thể xem điểm trung bình của mình và xem điểm từng bài tập.

- Người dùng khi nhấn tin đến người khác thì người kia sẽ hiện tin nhắn của người gửi tới.
- Tìm tài liệu từ tài liệu giảng viên của lớp đã tải tài liệu lên ,khi tìm kiếm nếu có tài liệu liên quan nó sẽ hiện hình ảnh tài liệu tên tài liệu và mục xem chi tiết và tải xuống của tài liệu ,không có kết quả tìm kiếm nó sẽ hiện thông báo không có tài liệu.
- Kiểm tra đạo văn : Bài tập sau khi được nộp thì khi giảng viên muốn kiểm tra đạo văn thì giảng viên sẽ sử dụng chức năng kiểm tra đạo văn và hiện phần trăm giống nhau nhiều nhất so với các tài liệu trước.
- Gom cụm dữ liệu : Để tối ưu hiệu suất cũng như tốc độ kiểm tra đạo văn, chức năng gom cụm dữ liệu sẽ tự động chạy cách mỗi hai tiếng một lần.

3. Lựa chọn công cụ cài đặt

Vì hệ thống được xây dựng trên internet nên em lựa chọn ứng dụng web bao gồm : Công nghệ ASP.NET MVC ,ngôn ngữ c#, ngôn ngữ golang ,hệ quản trị cơ sở dữ liệu SQL 2019 ,Công nghệ AJAX,Công nghệ runtime ...

4. Phạm vi đề tài

Nghiên cứu công nghệ runtime, cách thức hoạt động ,cách tạo bài tập tự luận và trắc nghiệm bằng file, thống kê điểm dựa trên bài tập đã có điểm để đảm bảo cho việc chấm điểm của sinh viên minh bạch. Cùng với đó là chức năng kiểm tra đạo văn bài tập tự luận giúp giảng viên kiểm tra bài tập và chấm điểm nhanh hơn. Bên cạnh đó nghiên cứu các phương pháp tổ chức hệ thống, đảm bảo hệ thống luôn trong trạng thái dễ phát triển và bảo trì về sau.

CHƯƠNG 1: CƠ SỞ LÝ THUYẾT

1.1. Công nghệ sử dụng

1.1.1. ASP.NET MVC

1.1.1.1. Mô hình MVC

Mô hình MVC (Model-View-Controller) một mô hình kiến trúc phần mềm phổ biến được sử dụng trong phát triển ứng dụng web và các ứng dụng máy tính khác. Nó tách biệt và tổ chức ứng dụng thành ba phần chính:

- Model: Đại diện dữ liệu và logic liên quan đến dữ liệu. Đảm nhận nhiệm vụ lưu trữ, truy xuất và xử lý dữ liệu.
- View: Đại diện cho giao diện người dùng. Hiển thị dữ liệu cho người dùng và nhận tương tác từ người dùng.
- Controller: Điều khiển và điều phối các tương tác giữa Model và View. Nhận tương tác từ người dùng thông qua View và cập nhật Model tương ứng.

Mô hình MVC giúp tách biệt logic xử lý dữ liệu, giao diện người dùng và điều khiển trong ứng dụng. Giúp tăng tính linh hoạt, dễ bảo trì và phát triển lại ứng dụng. Cho phép phân công công việc cho các nhóm phát triển khác nhau, cải thiện quy trình phát triển phần mềm.

1.1.1.2. ASP.NET MVC

ASP.NET MVC một framework phát triển ứng dụng web được xây dựng trên nền tảng .NET của Microsoft.

Trong ASP.NET MVC, ứng dụng được chia thành các phần chính:

- Model: Đại diện cho dữ liệu và logic liên quan đến dữ liệu. Model trong ASP.NET MVC thường là các lớp đại diện cho đối tượng dữ liệu, kết nối cơ sở dữ liệu hoặc các dịch vụ khác. Nó cung cấp các phương thức để truy xuất, xử lý và lưu trữ dữ liệu.

- View: Đại diện cho giao diện người dùng. View là các tệp HTML hoặc Razor chứa mã để hiển thị dữ liệu cho người dùng. Cho phép sử dụng các thẻ và cú pháp đặc biệt để trích xuất và hiển thị dữ liệu từ Model.
- Controller : Điều khiển và điều phối các tương tác giữa Model và View. Controller nhận các yêu cầu từ người dùng thông qua các tương tác trên giao diện người dùng và xử lý chúng. Truy xuất và cập nhật dữ liệu từ Model và sau đó chuyển dữ liệu cho View hiển thị cho người dùng.

ASP.NET MVC cung cấp một cách linh hoạt để phát triển ứng dụng web. Tách biệt logic xử lý, giao diện người dùng và dữ liệu, cho phép phân chia công việc cho các nhóm phát triển khác nhau và tăng tính tái sử dụng mã. Cung cấp nhiều tính năng như định tuyến, kiểm tra tính hợp lệ, và hỗ trợ cho AJAX.

1.1.2. Bootstrap

Một framework mã nguồn mở phổ biến được sử dụng để phát triển giao diện người dùng cho các trang web và ứng dụng web. Được phát triển bởi Twitter và được viết bằng HTML, CSS và JavaScript.

Bootstrap cung cấp một tập hợp các lớp CSS và các thành phần giao diện người dùng được thiết kế sẵn, giúp bạn xây dựng giao diện đáng tin cậy, linh hoạt và hấp dẫn mà không cần viết lại nhiều mã từ đầu.

Một số đặc điểm của Bootstrap :

- Hệ thống lưới: Bootstrap cung cấp một hệ thống lưới linh hoạt, cho phép bạn chia bố cục của trang thành các cột và hàng giúp tạo ra các giao diện đáp ứng, tự động điều chỉnh với các kích thước màn hình khác nhau.
- Bootstrap đi kèm với một loạt các thành phần UI như nút, biểu mẫu, thanh điều hướng, tiêu đề, hộp thông báo, bảng ... Các thành phần này được thiết kế sẵn và tương thích với nhau, giúp xây dựng giao diện nhất quán và chuyên nghiệp.
- Tùy chỉnh dễ dàng: Bootstrap cho phép tùy chỉnh giao diện bằng cách ghi đè CSS hoặc sử dụng Scss. Có thể điều chỉnh màu sắc, kích thước, kiểu chữ, hình

dạng và nhiều thuộc tính khác của các thành phần để phù hợp với yêu cầu của dự án.

- Hỗ trợ đa trình duyệt: Bootstrap tương thích với hầu hết các trình duyệt phổ biến như Chrome, Firefox, Safari, Edge và Internet Explorer. Đảm bảo rằng giao diện sẽ hoạt động trên các nền tảng và trình duyệt khác nhau.

Bootstrap giúp các developer có thể nhanh chóng xây dựng giao diện đáng tin cậy và tương thích trên nhiều thiết bị khác nhau, tiết kiệm thời gian và công sức trong việc phát triển giao diện người dùng cho ứng dụng web.

1.1.3. Signalr

Thư viện phát triển ứng dụng web thời gian thực của Microsoft. SignalR Runtime cung cấp môi trường cho việc xây dựng và triển khai ứng dụng sử dụng SignalR.

SignalR Runtime hỗ trợ các tính năng sau:

- Kết nối và duy trì liên lạc real-time: Cho phép xây dựng các ứng dụng web thời gian thực bằng cách cung cấp các phương pháp để thiết lập kết nối real-time giữa máy khách và máy chủ. Sử dụng các giao thức như WebSocket, Server-Sent Events hoặc Long Polling duy trì liên lạc liên tục giữa các phía.
- Broadcast và group: Cho phép gửi tin nhắn real-time từ máy chủ tới tất cả các kết nối máy khách hoặc một nhóm cụ thể. Điều này cho phép xây dựng ứng dụng với tính năng thông báo và cập nhật real-time.
- Giao tiếp dựa trên sự kiện: Cho phép giao tiếp real-time dựa trên sự kiện, máy khách và máy chủ gửi và lắng nghe các sự kiện với nhau.

SignalR Runtime tích hợp với các nền tảng web phổ biến như ASP.NET và ASP.NET Core, cung cấp các API và công cụ phát triển để xây dựng ứng dụng web real-time dễ dàng và hiệu quả.

Mô hình bình thường của web là client gửi request đến server, nhưng giờ có 1 cách thức nữa để tương tác: server “push” trực tiếp dữ liệu về client, theo thời gian thực. Ngoài ra, mô hình cũng giúp ta xây dựng ứng dụng bất đồng bộ lên đến hàng nghìn kết nối 1 lúc với máy chủ.

1.1.4. Spire.office

Kết hợp của API.NET Office doanh nghiệp do E-iceblue cung cấp. Bao gồm Spire.Doc, Spire.XLS, Spire.S Spreadsheet, Spire.Presentation, Spire.PDF, Spire.DataExport, Spire.OfficeViewer, Spire.PDFViewer, Spire.DocViewer, Spire.Barcode và Spire.Email. Spire.Office.

Spire.Office cho .NET có thể được liên kết với bất kỳ loại ứng dụng .NET 32 bit hoặc 64 bit bao gồm ASP.NET, Dịch vụ Web và WinForms cho .NET Framework phiên bản 2.0 đến 4.5. Spire.Office cũng hỗ trợ hoạt động trên .NET Core, .NET 5.0, Microsoft Azure, Mono Android và Xamarin.iOS.

Cho phép các developer mở, tạo, sửa đổi, chuyển đổi, in, xem các tài liệu MS Word, Excel, PowerPoint và PDF. Hơn nữa còn cho phép người dùng xuất dữ liệu sang các tệp phổ biến như MS Word / Excel / RTF / Access, PowerPoint, PDF, XPS, HTML, XML, Text, CSV, DBF, Clipboard, SYLK, PostScript, PCL,...

Ưu điểm:

- Hỗ trợ nhiều nền tảng.
- Hỗ trợ hầu hết các tính năng có trong Microsoft office.

Nhược điểm:

- Giới hạn số lượng trang cho phiên bản miễn phí.
- Phí bản quyền cao.

1.1.5. Cronjob

Một công cụ được sử dụng trong hệ điều hành để lập lịch và thực thi các tác vụ tự động theo một lịch trình cụ thể. Cronjob cho phép lập lịch cho các công việc để chạy theo một chu kỳ cố định, ví dụ: hàng ngày, hàng tuần, hàng tháng hoặc theo một định dạng thời gian tùy chỉnh.

Được sử dụng rộng rãi trong các hệ thống Unix và Unix-like (bao gồm Linux) để tự động hóa các tác vụ hàng ngày, như sao lưu dữ liệu, quét và cập nhật phần mềm, tạo báo cáo định kỳ, và nhiều công việc khác.

Một Cronjob bao gồm các thành phần sau:

- **Lịch trình cron:** Định rõ thời gian và ngày thực hiện công việc. Lịch trình cron được biểu diễn bằng một chuỗi các giá trị và ký tự đặc biệt như " * " (đại diện cho mọi giá trị), số (ví dụ: 0-59 cho phút, 0-23 cho giờ), hoặc các giá trị cụ thể (ví dụ: 0,15,30,45 cho phút là 0, 15, 30 và 45).
- **Công việc cần thực hiện:** Đó là lệnh hoặc tác vụ muốn thực hiện. Có thể là một lệnh dòng lệnh, một script, hoặc một chương trình.

Khi một Cronjob được thiết lập, hệ thống sẽ tự động chạy các công việc vào các thời điểm được chỉ định trong lịch trình cron. Kết quả của công việc thường được gửi qua email cho người dùng hoặc ghi vào tệp nhật ký.

Giúp tự động hóa các tác vụ định kỳ trong hệ thống, giúp tiết kiệm thời gian và công sức của người quản trị hệ thống.

1.1.6. Echo

Một framework của golang, giúp nhà phát triển xây dựng ứng dụng web và api. Được thiết kế với mục đích tối ưu hiệu suất, khả năng mở rộng, phát triển nhanh chóng và dễ dàng.

Một số đặc điểm chính của echo framework:

- **Router:** Định nghĩa các route và xử lý cho từng route, quyết định cách xử lý các yêu cầu HTTP từ client.
- **Handler:** Handler là một hàm hoặc phương thức xử lý yêu cầu HTTP. Định nghĩa handler để xử lý các route cụ thể. Handler nhận đầu vào là một context đại diện cho yêu cầu và phản hồi HTTP.
- **Context:** một đối tượng được tạo ra cho mỗi yêu cầu HTTP. Chứa thông tin về yêu cầu và phản hồi, cho phép truy cập vào các thông tin như tham số URL, header, body, và điều khiển quá trình xử lý.
- **Middleware:** Hàm trung gian được chạy trước hoặc sau khi xử lý yêu cầu HTTP. Cho phép thực hiện các tác vụ như xác thực, ghi log, nén dữ liệu, và xử lý lỗi trước khi yêu cầu đến handler chính.

- **Binding và Validation:** Cung cấp các công cụ để ràng buộc và xác thực dữ liệu đầu vào từ yêu cầu HTTP. Ràng buộc dữ liệu đầu vào từ query parameters, path parameters, form data, JSON, và nhiều loại dữ liệu khác.
- **Template:** Hỗ trợ việc sử dụng các template engine để tạo các trang HTML động. Sử dụng template engine như HTML/template, Jinja2, hoặc handlebars tạo các giao diện người dùng đẹp và tương tác.
- **HTTP Middleware:** Hỗ trợ sử dụng các middleware HTTP để cấu hình và tùy chỉnh các cài đặt HTTP, xử lý CORS, nén dữ liệu, cache control...
- **WebSocket:** Hỗ trợ việc sử dụng WebSocket để xây dựng các ứng dụng thời gian thực. Tạo các endpoint WebSocket để truyền dữ liệu theo thời gian thực giữa client và server.

Echo cung cấp nhiều tính năng và công cụ hỗ trợ để xây dựng ứng dụng web mạnh mẽ và linh hoạt trong Golang.

1.2. Ngôn ngữ lập trình

1.2.1. C#

Ngôn ngữ lập trình đa mục đích, hiện đại và mạnh mẽ, được phát triển bởi Microsoft. Một phần của .NET Framework và được sử dụng xây dựng ứng dụng Windows, ứng dụng web và ứng dụng di động.

Một số đặc điểm chính của C#:

- **Đơn giản và dễ học:** Được thiết kế với cú pháp gần với ngôn ngữ tự nhiên, dễ học và sử dụng cho những người mới học lập trình. Có cú pháp tương tự như các ngôn ngữ lập trình khác như C++ và Java.
- **Mạnh mẽ và linh hoạt:** Hỗ trợ các tính năng mạnh mẽ như quản lý bộ nhớ tự động, kiểu dữ liệu an toàn, quản lý ngoại lệ và lập trình hướng đối tượng. Hỗ trợ các tính năng tiên tiến như delegates, events, lambda expressions và LINQ (Language Integrated Query).

- Đa nền tảng: Không chỉ giới hạn trong việc phát triển ứng dụng trên hệ điều hành Windows. Với sự ra đời của .NET Core, có thể sử dụng C# để phát triển ứng dụng đa nền tảng trên Windows, macOS và Linux.
- Đa dạng ứng dụng: Được sử dụng để phát triển nhiều loại ứng dụng khác nhau như ứng dụng Windows Desktop, ứng dụng di động sử dụng Xamarin, ứng dụng web ASP.NET và ASP.NET Core, ứng dụng máy tính, game và nhiều ứng dụng khác.
- Ecosystem mạnh mẽ: Cung cấp một loạt các thư viện và công cụ hỗ trợ phát triển ứng dụng. Có thể sử dụng các thư viện sẵn có như Windows Forms, WPF, ASP.NET MVC, Entity Framework, và nhiều thư viện khác giúp việc phát triển ứng dụng trở nên dễ dàng hơn.

Với những ưu điểm và tính năng mạnh mẽ, C# đã trở thành một ngôn ngữ lập trình phổ biến trong cộng đồng phát triển phần mềm và được sử dụng rộng rãi trong các dự.

1.2.2. Javascript

JavaScript là một ngôn ngữ lập trình phía client, thông dịch và đa nền tảng. Được sử dụng chủ yếu để tạo các ứng dụng web tương tác, cung cấp khả năng tương tác và thay đổi nội dung của trang web mà không cần phải tải lại trang.

Một số đặc điểm chính của JavaScript:

- Ngôn ngữ phía client: Được thực thi trên trình duyệt của người dùng, không yêu cầu máy chủ để thực thi. Cho phép tương tác trực tiếp với giao diện người dùng và thay đổi nội dung của trang web.
- Đa nền tảng: JavaScript là ngôn ngữ đa nền tảng, chạy trên hầu hết các trình duyệt web phổ biến như Chrome, Firefox, Safari, và Edge. Ngoài ra, JavaScript cũng được sử dụng để phát triển ứng dụng di động thông qua các framework như React Native hoặc Ionic.
- Tương tác với HTML/CSS: Được sử dụng để tương tác với HTML và CSS để thay đổi, tạo ra hoặc xóa các phần tử, thay đổi kiểu dáng, thêm hiệu ứng và xử lý sự kiện trên trang web.

- **Cú pháp linh hoạt:** JavaScript có cú pháp dễ học và linh hoạt. Hỗ trợ nhiều kiểu dữ liệu như số, chuỗi, mảng và đối tượng. Cho phép xử lý điều kiện, vòng lặp, hàm và các phương thức xử lý sự kiện.

Hỗ trợ thư viện và framework mạnh mẽ: JavaScript có một hệ sinh thái phong phú với nhiều thư viện và framework như jQuery, React, Angular, Vue.js, và Node.js.... Giúp đơn giản hóa và tăng tốc quá trình phát triển ứng dụng JavaScript.

JavaScript đã trở thành một công nghệ quan trọng trong việc phát triển ứng dụng web và di động. Cung cấp khả năng tạo ra giao diện người dùng tương tác, xử lý dữ liệu và tương tác với các dịch vụ và API khác, làm cho trải nghiệm người dùng trên web trở nên linh hoạt.

1.2.3. Giới thiệu về HTML

Ngôn ngữ đánh dấu được sử dụng để tạo nội dung và cấu trúc trang web. Sử dụng các thẻ để định dạng và xác định các phần tử trên trang web.

Một số điểm cơ bản về HTML:

- **Cú pháp:** Được viết dưới dạng các thẻ mở và đóng, được bao quanh bởi các dấu ngoặc nhọn ("`<`" và "`>`"). Có các thẻ định dạng cơ bản như `<h1>`, `<p>`, `<a>`, ``...
- **Cấu trúc cơ bản:** Một trang HTML thông thường bao gồm các phần như tiêu đề (`<head>`), nội dung (`<body>`), các phần tử định dạng (`<header>`, `<nav>`, `<footer>`) và các phần tử chứa nội dung như đoạn văn bản, hình ảnh, liên kết và bảng.
- **Thuộc tính:** Các thẻ HTML có các thuộc tính để xác định các đặc tính của phần tử. Ví dụ, thuộc tính `src` của thẻ `` xác định nguồn ảnh, thuộc tính `href` của thẻ `<a>` để xác định liên kết.
- **CSS:** HTML thường được kết hợp với CSS để xác định giao diện và định dạng của trang web. Cho phép tạo các quy tắc định dạng, mà sau đó áp dụng cho các phần tử HTML.
- **Phiên bản:** Có nhiều phiên bản HTML như HTML4, XHTML và HTML5.

HTML là một phần quan trọng của việc xây dựng trang web. Cung cấp cấu trúc và định dạng cho nội dung, và cùng với CSS và JavaScript, tạo nên trải nghiệm đa dạng và tương tác trên web.

1.2.4. Giới thiệu về CSS

Một ngôn ngữ định dạng được sử dụng để xác định giao diện và định dạng của các trang web và tài liệu HTML. Cho phép điều chỉnh màu sắc, kiểu chữ, kích thước, khoảng cách và bố cục của các phần tử trên trang web.

Một số điểm cơ bản về CSS:

- **Cú pháp:** CSS sử dụng các quy tắc định dạng để áp dụng cho các phần tử HTML. Mỗi quy tắc bao gồm một bộ chọn (selector) và một khối khai báo (declaration block).
- **Bộ chọn (Selector):** Bộ chọn xác định các phần tử HTML muốn áp dụng định dạng. Có nhiều loại bộ chọn như bộ chọn phần tử (element selector), bộ chọn lớp (class selector), bộ chọn ID (ID selector), bộ chọn thuộc tính (attribute selector) và nhiều loại khác.
- **Thuộc tính và giá trị:** Mỗi quy tắc trong CSS bao gồm một hoặc nhiều cặp thuộc tính và giá trị. Thuộc tính xác định đặc tính cụ thể của phần tử, chẳng hạn như màu sắc, kích thước chữ, độ trong suốt,...
- **Kế thừa và ưu tiên:** CSS sử dụng cơ chế kế thừa để áp dụng các định dạng từ phần tử cha đến phần tử con. CSS cũng có quy tắc ưu tiên, cho phép ghi đè các định dạng được kế thừa hoặc xác định định dạng cụ thể cho một phần tử.
- **Tệp CSS ngoài:** CSS được xác định trong cùng tệp HTML hoặc trong các tệp CSS riêng biệt. Sử dụng tệp CSS ngoài giúp tổ chức mã CSS một cách tốt hơn và cho phép tái sử dụng mã CSS trên nhiều trang.

CSS là một phần quan trọng của phát triển web, cho phép tạo ra giao diện và định dạng tùy chỉnh cho trang web của mình. Một phần không thể thiếu trong việc tạo ra trải nghiệm người dùng hấp dẫn và trực quan trên web.

1.2.5. Golang

Một ngôn ngữ lập trình được phát triển bởi Google vào năm 2007. Go được thiết kế với mục tiêu tối ưu hóa hiệu suất và đơn giản hóa việc phát triển phần mềm.

Một số đặc điểm chính của Go:

- Cú pháp đơn giản: Go có một cú pháp rõ ràng, đơn giản và dễ đọc. Loại bỏ đi một số tính năng phức tạp của ngôn ngữ khác như lớp, kế thừa, nạp chồng phương thức, và quản lý bộ nhớ thủ công. Giúp làm giảm độ phức tạp của mã và tạo ra mã nguồn dễ bảo trì.
- Hiệu suất cao: Go được thiết kế để có hiệu suất cao. Sử dụng mô hình goroutine để xử lý đa luồng một cách hiệu quả và sử dụng bộ thu gom rác để tự động quản lý bộ nhớ. Giúp Go thích hợp cho việc xử lý đồng thời và xử lý các tác vụ có yêu cầu hiệu suất cao.
- Hỗ trợ đa nền tảng: Go có thể chạy trên nhiều hệ điều hành và kiến trúc khác nhau như Windows, macOS, Linux và các hệ điều hành nhúng. Một lựa chọn phù hợp cho việc phát triển ứng dụng di động, dịch vụ web và hệ thống phân tán.
- Hệ sinh thái công cụ và thư viện phong phú: Go đi kèm với một hệ sinh thái công cụ và thư viện phong phú, bao gồm trình biên dịch, trình quản lý gói, bộ kiểm tra tự động và các thư viện tiện ích. Hỗ trợ xây dựng ứng dụng web, phát triển game, xử lý dữ liệu và nhiều lĩnh vực khác...

Go đã trở thành một ngôn ngữ lập trình ngày càng phổ biến, được sử dụng trong nhiều dự án phát triển phần mềm từ những ứng dụng web đơn giản cho đến hệ thống phức tạp. Hiệu suất cao, cú pháp đơn giản và hệ sinh thái phong phú Go trở thành một lựa chọn hấp dẫn của các lập trình viên.

1.3. Cơ sở dữ liệu

1.3.1. Hệ quản trị cơ sở dữ liệu

Một hệ quản trị cơ sở dữ liệu quan hệ phát triển bởi Microsoft. Cung cấp một môi trường mạnh mẽ để quản lý, lưu trữ và truy vấn cơ sở dữ liệu.

Một số điểm cơ bản về SQL Server:

- Quản lý cơ sở dữ liệu: Cho phép tạo, cấu hình và quản lý cơ sở dữ liệu. Tạo các bảng, chỉ mục, ràng buộc và quan hệ giữa các bảng. SQL Server hỗ trợ nhiều tính năng như xác thực và phân quyền, sao lưu và phục hồi, và khả năng mở rộng để xử lý dữ liệu lớn.
- Ngôn ngữ truy vấn: Hỗ trợ ngôn ngữ truy vấn SQL (Structured Query Language) để truy vấn và thao tác dữ liệu trong cơ sở dữ liệu. SQL Server hỗ trợ một tập lệnh phong phú để truy vấn dữ liệu, thêm, sửa đổi và xóa dữ liệu từ các bảng, và thực hiện các thao tác khác như liên kết, nhóm và sắp xếp dữ liệu.
- Tính năng nâng cao: Cung cấp nhiều tính năng nâng cao để tăng cường hiệu suất và khả năng mở rộng của cơ sở dữ liệu như tạo chỉ mục, quản lý bộ nhớ đệm, tối ưu hóa truy vấn, cơ chế phân phối và khả năng lập lịch tự động.
- Tích hợp với công nghệ Microsoft: SQL Server tích hợp tốt với các công nghệ và dịch vụ khác của Microsoft. Có thể kết hợp với .NET Framework, Azure Cloud, Reporting Services, Integration Services và Analysis Services để cung cấp các giải pháp toàn diện cho việc quản lý và phân tích dữ liệu.
- Hỗ trợ đa nền tảng: Đã có phiên bản SQL Server cho Linux, mở rộng sự hỗ trợ của nó trên các nền tảng khác ngoài Windows.

SQL Server là một hệ quản trị cơ sở dữ liệu phổ biến và được sử dụng rộng rãi trong các ứng dụng doanh nghiệp và web. Cung cấp các tính năng mạnh mẽ và linh hoạt để quản lý và truy vấn cơ sở dữ liệu.

1.3.2. Cơ sở dữ liệu

Một tập hợp có tổ chức các dữ liệu liên quan đến nhau và được lưu trữ trong một hệ thống quản lý cơ sở dữ liệu. Cơ sở dữ liệu được sử dụng để lưu trữ, quản lý và tương tác với dữ liệu trong một hệ thống thông tin.

Cơ sở dữ liệu có thể được tổ chức theo nhiều mô hình khác nhau, nhưng mô hình quan hệ mô hình phổ biến nhất trong các hệ thống cơ sở dữ liệu hiện đại. Trong mô hình quan hệ, dữ liệu được tổ chức thành các bảng có mối quan hệ với nhau thông qua các khóa.

Cơ sở dữ liệu bao gồm:

- **Bảng:** Là một đơn vị cơ bản để lưu trữ dữ liệu trong cơ sở dữ liệu quan hệ. Bảng bao gồm các hàng và cột và mỗi hàng đại diện cho một bản ghi cụ thể.
- **Cột:** Đại diện cho một thuộc tính cụ thể trong bảng. Mỗi cột có một kiểu dữ liệu như số nguyên, chuỗi, ngày tháng ...
- **Hàng:** Đại diện cho một bản ghi cụ thể trong bảng. Mỗi hàng chứa dữ liệu cho tất cả các cột tương ứng trong bảng.
- **Khóa :** Được sử dụng để định danh duy nhất các bản ghi trong bảng. Khóa chính là một khóa duy nhất cho mỗi bản ghi, trong khi khóa ngoại là một liên kết với khóa chính của bảng khác.
- **Truy vấn:** Là một câu lệnh hoặc một tập hợp các câu lệnh được sử dụng để truy xuất, cập nhật, thêm, xóa dữ liệu trong cơ sở dữ liệu. SQL ngôn ngữ phổ biến được sử dụng để tương tác với cơ sở dữ liệu.

Cơ sở dữ liệu đóng vai trò quan trọng trong các ứng dụng và hệ thống thông tin hiện đại, từ ứng dụng di động đến hệ thống quản lý khách hàng, thương mại điện tử, và nhiều lĩnh vực khác.

1.4. Thuật toán gom cụm dữ liệu

Thuật toán phân cụm dữ liệu (Clustering Algorithm) một phương pháp trong lĩnh vực học máy và khám phá dữ liệu để nhóm các đối tượng dữ liệu có tính chất tương tự vào các nhóm (cụm) khác nhau dựa trên một số tiêu chí nhất định. Mục tiêu tạo ra các nhóm nội bộ có tính tương đồng cao và tính khác biệt giữa các nhóm.

1.4.1. K-means

K-means một thuật toán phân cụm dữ liệu không giám sát, phổ biến trong lĩnh vực trích xuất thông tin và khai phá dữ liệu. Thuật toán giúp nhóm các điểm dữ liệu vào các cụm dựa trên đặc trưng tương tự nhau.

Các bước của thuật toán kmeans:

- **Bước 1 :** Chọn số K cụm dữ liệu muốn phân chia.
- **Bước 2 :** Chọn ngẫu nhiên K điểm làm điểm trung tâm ban đầu của các cụm.

- Bước 3 :Gán mỗi điểm dữ liệu vào cụm có điểm trung tâm gần nhất.
- Bước 4 :Tính toán lại các điểm trung tâm của các cụm dựa trên các điểm dữ liệu đã được gán vào từng cụm.
- Bước 5 :Lặp lại bước 3 và 4 cho đến khi các điểm trung tâm không thay đổi hoặc đạt đến số lần lặp tối đa.

Kết quả của thuật toán K-means là một tập hợp K cụm, trong đó mỗi điểm dữ liệu được gán vào một cụm duy nhất. Mỗi cụm được đại diện bởi một điểm trung tâm, được tính bằng trung bình của các điểm dữ liệu trong cụm.

Ưu điểm : Dễ triển khai và hiệu quả đối với dữ liệu có số lượng lớn và không quá phức tạp.

Nhược điểm : Kết quả của K-means phụ thuộc vào số cụm K ban đầu, không phù hợp với dữ liệu có cấu trúc phức tạp hoặc kích thước cụm không đồng đều. Với các trung tâm ban đầu khác nhau có thể dẫn đến kết quả khác nhau mỗi lần chạy.

Độ phức tạp của kmeans : $O(t * n * k * d)$

Với số lượng điểm dữ liệu (n), số lượng cụm (k), số lần lặp (t) và số chiều của không gian dữ liệu (d)

1.4.2. Agglomerative Hierarchical Clustering

Agglomerative Hierarchical Clustering một thuật toán gom cụm phổ biến trong machine learning và phân tích dữ liệu. Mỗi điểm dữ liệu ban đầu được xem là một cụm riêng biệt sau đó các cụm được gom lại dựa trên mức độ tương đồng

Thuật toán bắt đầu bằng việc xem xét mỗi điểm dữ liệu như một cụm độc lập. Các cặp cụm gần nhất được tìm thấy dựa trên một phép đo tương đồng như khoảng cách Euclidean hoặc cosin và được gom lại thành một cụm mới. Tiếp tục đến khi chỉ còn lại một cụm duy nhất chứa tất cả các điểm dữ liệu.

Agglomerative Hierarchical Clustering có hai cách tiếp cận chính để đo độ tương đồng giữa các cụm:

- Single-linkage clustering: khoảng cách giữa hai cụm được tính bằng khoảng cách nhỏ nhất giữa hai điểm thuộc hai cụm tương ứng. Cách tiếp cận đơn giản và nhạy cảm với các cụm dài và xoắn.
- Complete-linkage clustering: khoảng cách giữa hai cụm được tính bằng khoảng cách lớn nhất giữa hai điểm thuộc hai cụm tương ứng. Giúp giảm nhạy cảm với các cụm dài xoắn và tạo ra các cụm có kích thước tương đối đồng đều.

Kết quả của thuật toán là cây phân cấp, biểu diễn cấu trúc phân cụm từ cụm lớn nhất chứa tất cả các điểm dữ liệu đến các cụm nhỏ.

Ưu điểm:

- Không cần xác định trước số lượng cụm
- Cấu trúc phân cấp: cung cấp một cây phân cấp biểu diễn cấu trúc phân cụm từ cụm lớn nhất chứa tất cả các điểm dữ liệu đến các cụm nhỏ cung cấp cái nhìn toàn diện mối quan hệ giữa các cụm dữ liệu.
- Tích hợp linh hoạt: áp dụng cho nhiều loại dữ liệu và sử dụng các phép đo tương đồng linh hoạt để xác định mức độ tương đồng giữa các cụm.

Nhược điểm:

- Độ phức tạp tính toán cao: có độ phức tạp tính toán cao đặc biệt là khi số lượng điểm dữ liệu lớn.
- Nhạy cảm với nhiễu và đặc trưng không quan trọng
- Không thích hợp cho dữ liệu lớn: Với tập dữ liệu lớn gặp khó khăn trong việc xử lý và tiêu tốn nhiều tài nguyên tính toán.

1.4.2. Density-Based Spatial Clustering of Applications with Noise

Một thuật toán gom cụm dựa trên mật độ trong không gian dữ liệu. Có khả năng phát hiện cụm có hình dạng và kích thước không đều, cùng với khả năng xác định và loại bỏ điểm nhiễu.

Cách thức hoạt động :

- Xác định tham số: xác định độ lớn của vùng xung quanh mỗi điểm dữ liệu và số lượng điểm dữ liệu tối thiểu trong mỗi vùng và xem là một cụm.
- Xác định điểm cốt lõi: Một điểm dữ liệu có ít số lượng điểm dữ liệu tối thiểu trong mỗi vùng nằm trong vùng xung quanh của nó được xem là điểm cốt lõi và được xác định bởi epsilon.
- Xác định cụm: Một cụm bắt đầu từ một điểm cốt lõi và bao gồm tất cả các điểm dữ liệu có thể tiếp cận được từ điểm cốt lõi qua các kết nối mật độ. Điểm tiếp cận là các điểm dữ liệu có thể truy cập thông qua một loạt các điểm cốt lõi.
- Xác định điểm nhiễu: không phải là điểm cốt lõi và không thuộc bất kỳ cụm nào.

Ưu điểm:

- Khả năng phát hiện cụm có hình dạng và kích thước không đều
- Xác định và loại bỏ điểm nhiễu
- Không yêu cầu xác định trước số lượng cụm

Nhược điểm:

- Độ phức tạp tính toán cao: có độ phức tạp tính toán cao đặc biệt là khi số lượng điểm dữ liệu lớn.
- Nhạy cảm với tham số: yêu cầu xác định đúng các tham số như epsilon và số lượng điểm dữ liệu tối thiểu trong mỗi vùng để cho kết quả gom cụm tốt. Lựa chọn sai epsilon dẫn đến kết quả không chính xác.
- Không hiệu quả đối với dữ liệu có mật độ không đồng nhất: khi các cụm có mật độ khác nhau không hoạt động tốt.

1.4.2. Mean-Shift

Một thuật toán gom cụm không có giám sát trong lĩnh vực machine learning. Được sử dụng để tìm và phân cụm các điểm dữ liệu dựa trên mật độ trong không gian đặc trưng.

Cách thức hoạt động:

- Bước 1: Khởi tạo các điểm trung tâm: Mỗi điểm dữ liệu được chọn ngẫu nhiên làm một điểm trung tâm ban đầu.
- Bước 2: Tính toán vector shift: Với mỗi điểm trung tâm, lấy trung bình các điểm dữ liệu trong vùng xung quanh có bán kính được xác định trước ta được vector shift.
- Bước 3: Cập nhật điểm trung tâm: Di chuyển các điểm trung tâm theo vector shift tính toán.
- Bước 4: Lặp lại bước 2 và 3 đến khi các điểm trung tâm không còn thay đổi nhiều hoặc đạt tới một tiêu chí dừng nhất định.
- Bước 5: Gán các điểm dữ liệu vào các cụm: Mỗi điểm dữ liệu được gán vào cụm tương ứng với điểm trung tâm nó hội tụ đến.

Kết quả của thuật toán là các cụm được xác định dựa trên mật độ của dữ liệu. Các điểm dữ liệu có mật độ cao hơn sẽ hội tụ về các điểm trung tâm, trong khi các điểm dữ liệu ở các vùng thưa thớt sẽ nằm ở các điểm trung tâm khác.

Ưu điểm:

- Khả năng phát hiện cụm có hình dạng không đều: Khả năng tìm ra các cụm có hình dạng và kích thước không đều.
- Tự động xác định số lượng cụm
- Không yêu cầu giả định về phân phối dữ liệu: Không giả định rằng dữ liệu phân phối theo một phân phối cụ thể, phù hợp cho nhiều loại dữ liệu.

Nhược điểm

- Độ phức tạp tính toán cao: có độ phức tạp tính toán cao đặc biệt là khi số lượng điểm dữ liệu lớn
- Nhạy cảm với tham số: Việc lựa chọn bán kính ảnh hưởng đến kết quả của thuật toán.

Với các ưu điểm và nhược điểm trên thì việc chọn kmeans sẽ phù hợp với bài toán đang triển khai.

1.5. Thuật toán áp dụng

1.5.1. Thuật toán Kmeans

1.5.1.1 Triển khai thuật toán

Input: Danh sách tài liệu, k số lượng cụm cần chia

Output: k cụm danh sách tài liệu

Ví dụ :

Input:

- Danh sách tài liệu :

["C (ngôn ngữ lập trình) – Wikipedia tiếng Việt.",

"C cơ bản: Giới thiệu ngôn ngữ C - DevIOT.",

"Ngôn ngữ lập trình C là gì? Tìm hiểu về ngôn ngữ lập trình C.",

"Khi nào nên dùng Golang? Nó dùng tốt trong trường hợp nào?.",

"Golang là gì? Backend Developer có nên học Golang?.",

"Golang là gì? Vì sao nên sử dụng ngôn ngữ Golang.",

"Python (ngôn ngữ lập trình) – Wikipedia tiếng Việt.",

"Python Là Gì? Các Bước Tự Học Lập Trình Python - TopDev.",

"Python là gì? - Giải thích về ngôn ngữ Python - Amazon AWS."]

- K =3

Output :

{"list": {

"1": ["Python Là Gì? Các Bước Tự Học Lập Trình Python - TopDev.",

"Python (ngôn ngữ lập trình) – Wikipedia tiếng Việt."],

"2": ["C cơ bản: Giới thiệu ngôn ngữ C - DevIOT."],

"3": ["Golang là gì? Vì sao nên sử dụng ngôn ngữ Golang.",

"Golang là gì? Backend Developer có nên học Golang?.",
 "Python là gì? - Giải thích về ngôn ngữ Python - Amazon AWS.",
 "C (ngôn ngữ lập trình) – Wikipedia tiếng Việt.",
 "Khi nào nên dùng Golang? Nó dùng tốt trong trường hợp nào?.",
 "Ngôn ngữ lập trình C là gì? Tìm hiểu về ngôn ngữ lập trình C."] },
 "point": { "1": { "x": 6.7105263157894735,"y": 1.2310604636591485 },
 "2": { "x": 2.227272727272727,"y": 1.1820454545454544 },
 "3": { "x": 4.6,"y": 1.4847250000000005 } } }

Thuật toán :

Bước 1 : Lấy ra danh sách tài liệu trong hệ thống. Danh sách tài liệu được lưu trong bảng dưới đây

STT	Tên thuộc tính	Kiểu dữ liệu	Mô tả
1	listtailieu	Array	Danh sách tài liệu trong hệ thống

Bảng 1. 1 Bảng danh sách tài liệu

Bước 2 : Tách các từ trong danh sách tài liệu. Với từ được lưu trong bảng dưới đây.

STT	Tên thuộc tính	Kiểu dữ liệu	Mô tả
1	listtu	Array	Danh sách từ

Bảng 1. 2 Bảng danh sách từ

Bước 3: biến các từ trong tài liệu thành các điểm có tọa độ x y dựa vào chiều dài và giá trị byte. Với danh sách tọa độ của từ được lưu trong bảng dưới đây

STT	Tên thuộc tính	Kiểu dữ liệu	Mô tả
1	Listpoint	array	Danh sách tọa độ từ

Bảng 1. 3 Bảng danh sách cụm từ

Bước 4 : chọn k điểm từ listpoint. Với k điểm lưu trong bảng dưới đây

STT	Tên thuộc tính	Kiểu dữ liệu	Mô tả
1	listpointcenter	Array	Danh sách điểm trung tâm

Bảng 1. 4 Bảng danh sách điểm trung tâm

Bước 5 : Tính khoảng cách và phân chia các điểm vào k cụm. Với mỗi cụm điểm lưu trong bảng dưới đây

STT	Tên thuộc tính	Kiểu dữ liệu	Mô tả
1	Pointcenter	Array	Điểm trung tâm
2	Listpoint	Array	Danh sách điểm

Bảng 1. 5 Bảng danh sách cụm điểm

Bước 6 : Tính toán lại các điểm trung tâm và lặp lại bước 5 cho đến khi các điểm trung tâm không thay đổi.

Bước 7 : Sau khi thực hiện các bước trên ta sẽ có được danh sách k cụm dữ liệu và các điểm trung tâm.

1.5.2. Thuật toán kiểm tra đạo văn

1.5.2.1. Thuật toán linear search

Thuật toán tìm kiếm một giá trị cụ thể trong một mảng, hoạt động bằng cách duyệt qua từng phần tử trong danh sách cho đến khi tìm thấy phần tử cần tìm trong danh sách.

Các bước của thuật toán :

- Bước 1: Duyệt từng phần tử của mảng
- Bước 2: So sánh phần tử với giá trị cần tìm.
- Bước 3: Nếu giá trị phần tử bằng với giá trị cần tìm, quá trình tìm kiếm kết thúc và trả về vị trí của phần tử cần tìm trong mảng. Nếu phần tử hiện tại không trùng khớp với giá trị cần tìm, di chuyển đến phần tử tiếp theo trong mảng. Lặp lại các bước 1, 2, 3 cho đến khi tìm thấy phần tử hoặc duyệt qua toàn bộ danh sách.
- Bước 4: Nếu đã duyệt qua toàn bộ mảng mà không tìm thấy phần tử, thuật toán trả về một giá trị không hợp lệ để biểu thị rằng phần tử không tồn tại trong mảng.

Độ phức tạp thời gian của thuật toán tìm kiếm tuyến tính là $O(n)$, với n là kích thước của danh sách.

1.5.2.2 Triển khai thuật toán

Input: Tài liệu gốc, tài liệu cần so sánh

Output: Phần trăm đạo văn của tài liệu cần so sánh so với tài liệu gốc.

Thuật toán :

Bước 1 : Lấy ra danh sách các câu có trong tài liệu gốc và tài liệu cần so sánh dựa vào dấu “.” cuối câu. Với mỗi tài liệu được lưu trong bảng dưới đây

STT	Tên thuộc tính	Kiểu dữ liệu	Mô tả
1	listCau	Array	Danh sách các câu có trong tài liệu

Bảng 1. 6 Bảng danh sách câu

Bước 2 :Lấy câu của tài liệu cần so sánh so với từng câu của tài liệu gốc. Với mỗi câu cần so sánh được lưu trong bảng dưới đây.

STT	Tên thuộc tính	Kiểu dữ liệu	Mô tả
1	cau	string	Câu cần so sánh
2	Phantram	float	Phần trăm sau khi so sánh

Bảng 1. 7 Bảng phần trăm câu

Bước 3 : Gọi A1 là câu của tài liệu cần so sánh, B1 là câu của tài liệu gốc. Lấy ra các cụm từ trong A1 có trong B1. Với mỗi câu A1 các cụm được lưu vào bảng dưới đây.

STT	Tên thuộc tính	Kiểu dữ liệu	Mô tả
1	listCumtu	Array	Danh sách các cụm từ có trong câu A1 có trong B1

Bảng 1. 8 Bảng danh sách cụm từ

Bước 4 : Tính phần trăm độ giống nhau của từng câu dựa vào công thức

$$P[i] = \frac{\alpha[1] + \alpha[2] + \alpha[3] + \dots + \alpha[i]}{\beta} \times 100$$

- $\alpha[1], \alpha[2], \alpha[3], \dots$ số lượng từ trong các cụm từ
- β là số lượng từ trong câu A1.

Bước 5 : Thực hiện tương tự A1 với các câu của tài liệu gốc và lưu phần trăm vào bảng dưới đây :

STT	Tên thuộc tính	Kiểu dữ liệu	Mô tả
1	listphamtram	Array	Danh sách phần trăm giống nhau của A1 so với từng câu của tài liệu gốc

Bảng 1. 9 Bảng danh sách phần trăm

Bước 6: Sau khi so sánh câu của tài liệu cần so sánh so với từng câu của tài liệu gốc ta được 1 danh sách phần trăm giống nhau của câu của tài liệu cần so sánh so với từng câu của tài liệu gốc. Lấy ra phần tử lớn nhất trong danh sách phần trăm giống nhau. thực hiện các bước tương tự cho từng câu.

Bước 7 : Sau khi thực hiện xong các bước trên ta được danh sách phần trăm sau so sánh của từng câu. Thực hiện tính toán phần trăm đạo văn của tài liệu theo công thức sau :

$$\text{Percent} = \frac{\rho[1] + \rho[2] + \rho[3] + \dots + \rho[i]}{\alpha}$$

- $\rho[i]$ là phần trăm sau khi so sánh của từng câu.
- α là số lượng câu của tài liệu cần so sánh.

CHƯƠNG 2: PHÂN TÍCH HỆ THỐNG

2.1. Mô tả bài toán

LMS là phần mềm giúp tạo tài liệu, quản trị, báo cáo, theo dõi, phân phối và tự động hóa các chương trình đào tạo hoặc học tập và phát triển

LMS gồm các chức năng:

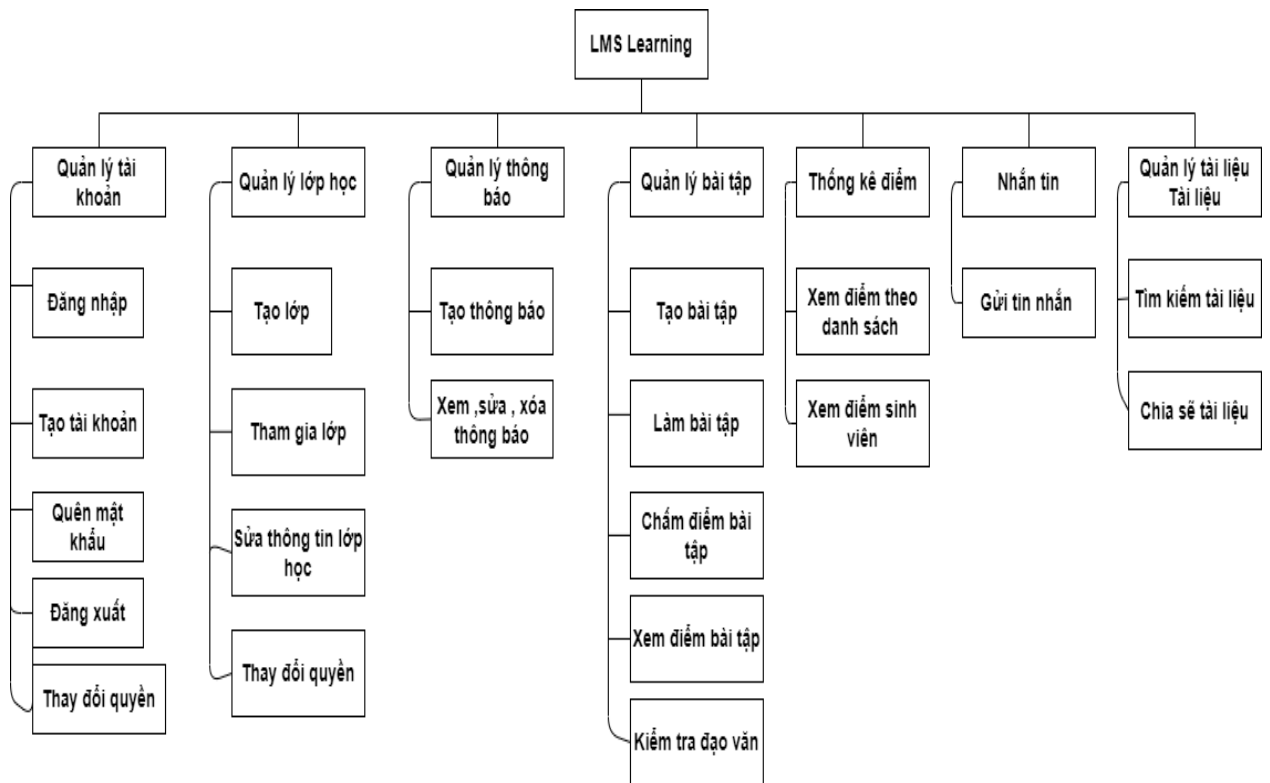
- Quản lý tài khoản : khi người dùng sử dụng hệ thống trước hết cần phải đăng nhập hệ thống. Nếu người dùng không có tài khoản thì cần tạo tài khoản mới với các thông tin cần có: họ, tên, tên tài khoản ,email ,mật khẩu. Sau khi điền đầy đủ thông tin thì người dùng cần nhập mã xác nhận được gửi tới tài khoản email đã đăng ký. Tài khoản sau khi đang kí chỉ có quyền sinh viên nên người dùng muốn tài khoản có các chức năng của giáo viên thì admin sẽ là người thay đổi quyền. Khi đăng nhập nếu người dùng quên mật khẩu có thể thực hiện chức năng quên mật khẩu. Chức năng cần có tên đăng nhập và email của người dùng. Sau khi hoàn tất các bước thì người dùng sẽ nhập mã xác nhận email. Nếu mã xác nhận đúng thì lúc này người dùng mới điền được mật khẩu mới của tài khoản. Khi người dùng sử dụng xong thì đăng xuất khỏi hệ thống.
- Quản lý lớp học :có các chức năng như hiện danh sách lớp học đã tham gia ,tạo lớp học mới ,tham gia lớp học. Đối với chức năng tham gia lớp học thì người dùng cần nhập mã lớp học do giảng viên lớp học cung cấp. Một lớp học có thể có nhiều giáo viên nên khi người dùng tham gia lớp học chỉ có quyền là sinh viên nên khi muốn thay đổi quyền thì admin sẽ làm các chức năng như thay đổi quyền lớp học.
- Quản lý thông báo : Có các chức năng như đăng thông báo, xem sửa xóa thông báo ,hiện thông báo mới. Với chức năng sửa xóa thông báo thì chỉ người đăng thông báo mới được quyền sửa dụng. Sau khi đăng thông báo thì những người trong lớp học đó sẽ hiện mail thông báo đến tài khoản. Khi người dùng đang trong trang thông báo thì khi có người đăng thông báo thì người dùng sẽ hiện thông báo để reload lại trang.

- Quản lý bài tập :có các chức năng như đăng bài tập sửa xóa bài tập ,chấm điểm bài tập ,nộp bài tập ,hiện bài tập mới.Với chức năng đăng bài tập thì có chỉ người giảng viên lớp đó mới được sử dụng. Bài tập có hai dạng đó là bài tự luận và bài trắc nghiệm. Bài trắc nghiệm thì được thêm bằng cách thêm file câu hỏi theo định dạng chương trình. Giảng viên có thể sửa xóa thông tin bài tập. Với chức năng chấm điểm thì với bài tập trắc nghiệm thì khi làm bài xong thì hệ thống tự chấm điểm ,còn với bài tập tự luận thì giảng viên dựa theo bài tập sinh viên nộp để chấm điểm. Thang điểm bài tập từ 0 đến 100 điểm. Giảng viên có thể kiểm tra đạo văn đối với bài tập tự luận, bài tập nộp đầu tiên trong lớp sẽ có phần trăm đạo văn là 0 sau đó thì tài liệu sẽ kiểm tra đạo văn so với tài liệu đã kiểm tra đạo văn. Kiểm tra đạo văn sẽ kiểm tra từ tài liệu nộp sớm nhất. Giảng viên có thể xem điểm của sinh viên theo bài tập. Khi giảng viên đăng bài tập thì hệ thống sẽ gửi email thông báo có bài tập cho sinh viên. Nếu sinh viên đang ở trong trang bài tập thì sẽ hiện thông báo reload lại trang. Sinh viên có thể xem thông tin bài tập đã nộp ,chưa nộp, bài tập nộp muộn.
- Thống kê điểm :có các chức năng xem điểm theo danh sách ,xem điểm sinh viên . Giảng viên có thể xem thông bài tập của sinh viên :như thông tin bài nộp ,bài chưa nộp bài nộp muộn và điểm trung bình theo danh sách sinh viên lớp. Sinh viên có thể xem các thông tin điểm trung bình của mình ,điểm của từng bài tập.
- Nhắn tin :có các chức năng gửi tin nhắn hiện thông báo có tin nhắn đến. Với chức năng gửi tin nhắn :người dùng gửi tin nhắn thì người nhận tin nhắn sẽ tin nhắn đến.Với chức năng hiện thông báo tin nhắn đến thì người dùng ở trong trang thì sẽ hiện thông báo có tin nhắn từ người khác gửi tới.
- Quản lý Tài liệu : có chức năng tìm tài liệu. Người dùng nhập thông tin muốn tìm kiếm thì sẽ hiện thông tin tài liệu người dùng muốn tìm. Người dùng có thể tải tài liệu xuống ,xem chi tiết của tài liệu.
- Kiểm tra đạo văn : Giảng viên sẽ có 3 người dữ liệu để kiểm tra :kiểm tra với các bài khác cùng bài tập, kiểm tra với nguồn dữ liệu của lớp học ,kiểm tra với tất cả dữ liệu của hệ thống. Đối với kiểm tra nguồn dữ liệu của hệ thống thì sẽ hệ

thông sẽ tự động tạo job và sẽ chờ batch processing chạy bằng golang sử lý vào 1 h hằng ngày (thời gian do người quản lý hệ thống tạo) để dữ lý dữ liệu.

- Gom cụm dữ liệu :hệ thống sẽ tự động phân cụm dữ liệu cách 2 h sẽ chạy một lần (thời gian do người quản lý hệ thống tạo).

2.2. Sơ đồ phân cấp chức năng BFD



Hình 2. 1 Sơ đồ phân cấp chức năng

2.3. Bảng phân tích xác định tiến trình, tác nhân và hồ sơ

2.3.1. Xác định tiến trình

Động từ + bổ ngữ	Danh từ	Nhận xét
Quản lý tài khoản	Sinh viên Giảng viên Admin Thông tin tài khoản Thông tin yêu cầu Thông tin phản hồi	Tác nhân Tác nhân Tác nhân Hồ sơ dữ liệu Hồ sơ dữ liệu Hồ sơ dữ liệu

Quản lý lớp học	Sinh viên Giảng viên Admin Thông tin lớp học Thông tin yêu cầu Thông tin phản hồi	Tác nhân Tác nhân Tác nhân Hồ sơ dữ liệu Hồ sơ dữ liệu Hồ sơ dữ liệu
Quản lý thông báo	Sinh viên Giảng viên Thông tin yêu cầu Thông tin phản hồi Thông tin	Tác nhân Tác nhân Hồ sơ dữ liệu Hồ sơ dữ liệu Hồ sơ dữ liệu
Quản lý bài tập	Sinh viên Giảng viên thông tin bài tập Thông tin yêu cầu Thông tin phản hồi Bài nộp Điểm Phần trăm đạo văn	Tác nhân Tác nhân Hồ sơ dữ liệu Hồ sơ dữ liệu Hồ sơ dữ liệu Hồ sơ dữ liệu Hồ sơ dữ liệu Hồ sơ dữ liệu
Thống kê điểm	Sinh viên Giảng viên Thông tin yêu cầu Điểm	Tác nhân Tác nhân Hồ sơ dữ liệu Hồ sơ dữ liệu
Nhắn tin	Sinh viên Giảng viên Tin nhắn đến Tin nhắn đi Thông tin yêu cầu Thông tin phản hồi	Tác nhân Tác nhân Hồ sơ dữ liệu Hồ sơ dữ liệu Hồ sơ dữ liệu Hồ sơ dữ liệu

	Thông tin	Hồ sơ dữ liệu
Tài liệu	Sinh viên Giảng viên Thông tin yêu cầu Tài liệu	Tác nhân Tác nhân Hồ sơ dữ liệu Hồ sơ dữ liệu

Bảng 2. 1 Bảng xác định tiến trình

2.3.2. Tác nhân và hồ sơ

a). Tác nhân

Sinh viên, giảng viên, admin

b). Hồ sơ dữ liệu

Tin nhắn đến ,tin nhắn đi

Thông tin tài khoản, thông tin yêu cầu, thông tin phản hồi

Thông tin

Tài liệu

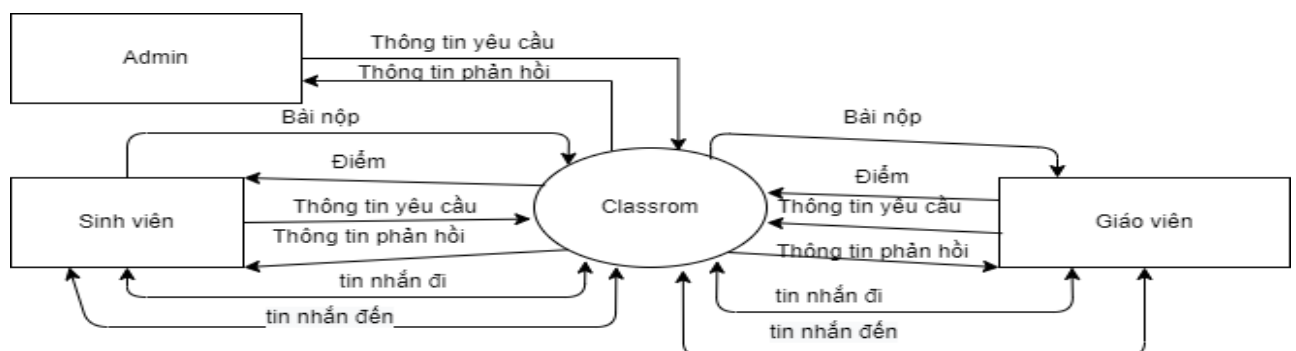
Bài nộp, điểm

Phản trảm đạo văn

Thông tin lớp học

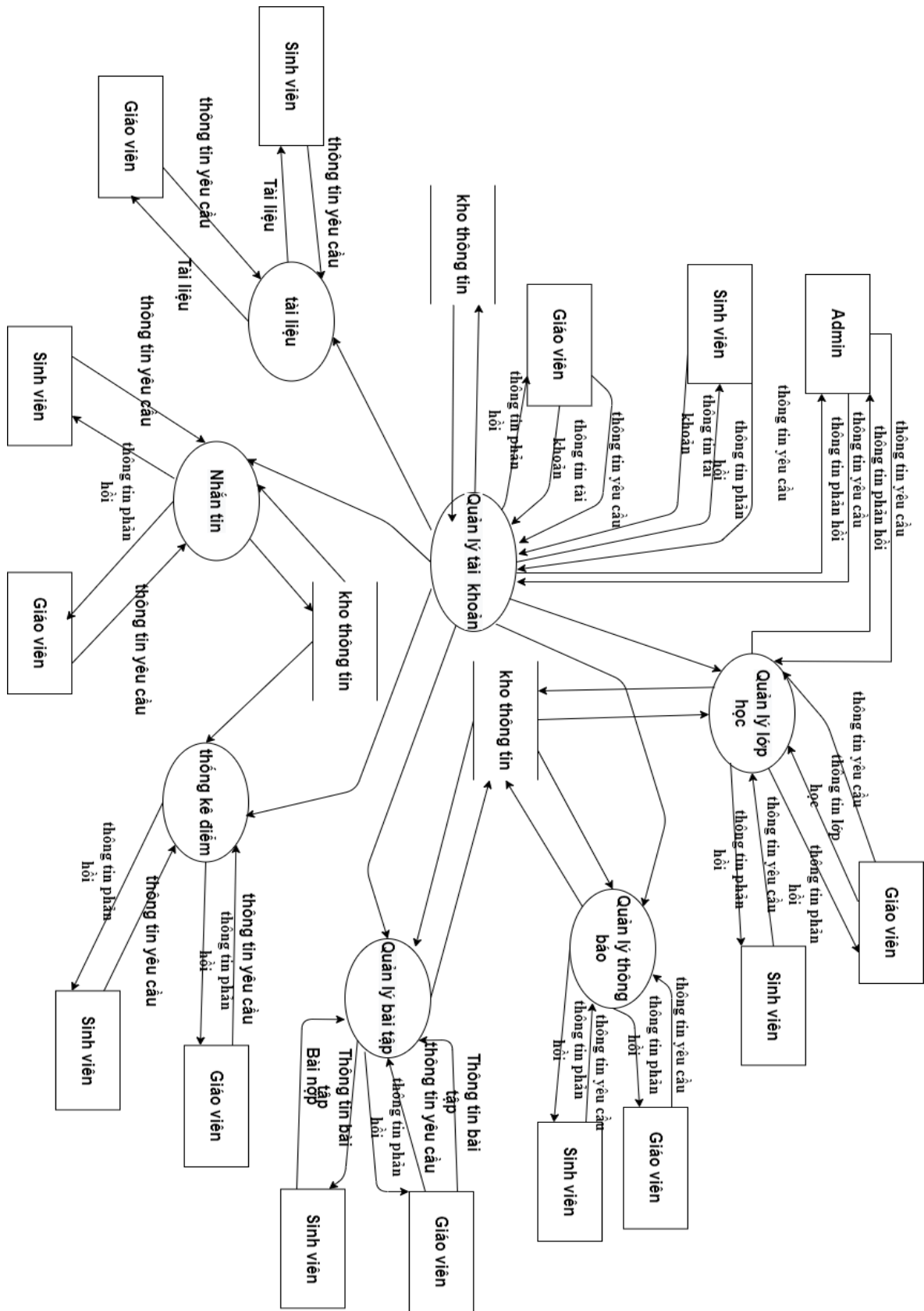
2.4. Biểu đồ luồng dữ liệu (mức ngữ cảnh, mức đỉnh, mức dưới đỉnh)

2.4.1. DFD mức ngữ cảnh



Hình 2. 2 DFD mức ngữ cảnh.

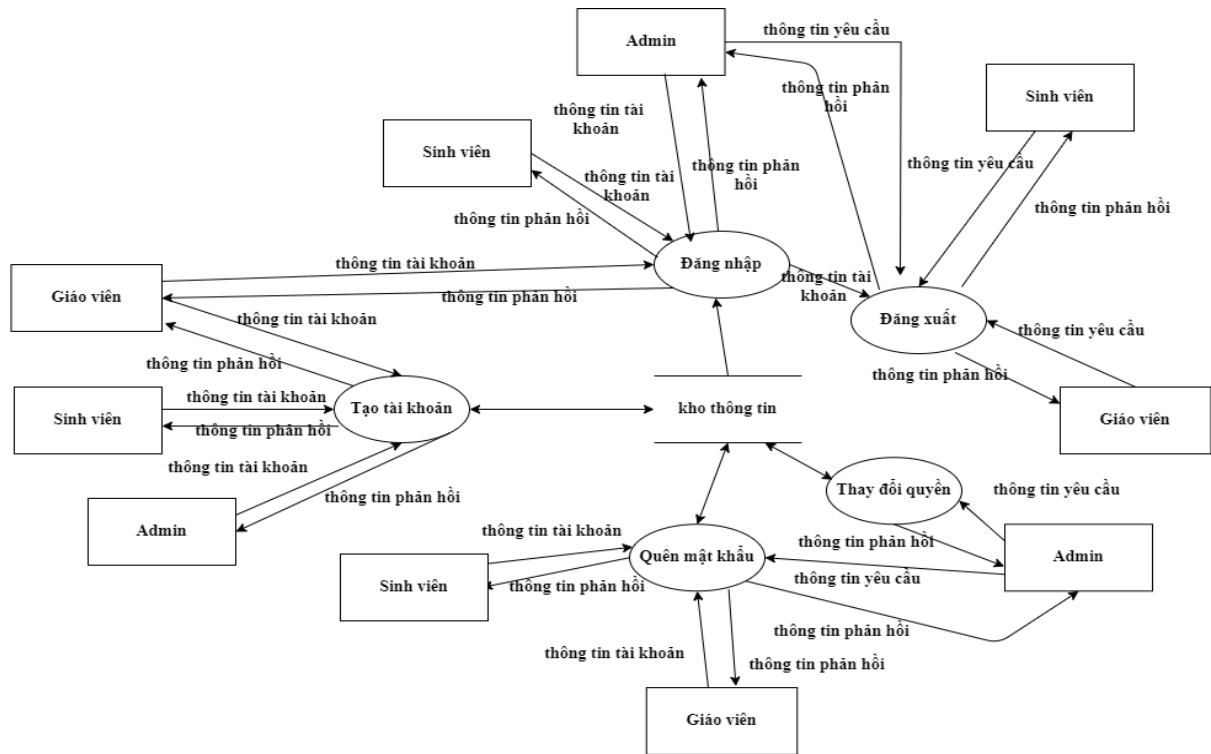
2.4.2. DFD mức đỉnh



Hình 2. 3 DFD mức đỉnh.

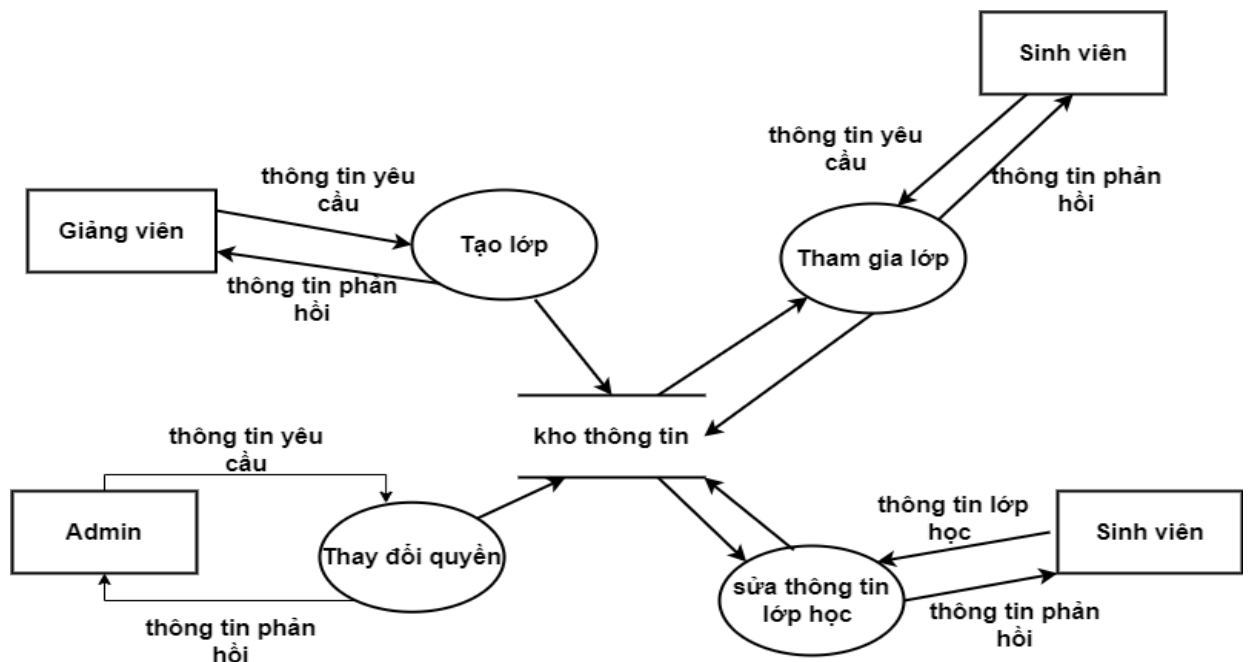
2.4.3. DFD mức dưới đỉnh

➤ Quản lý tài khoản



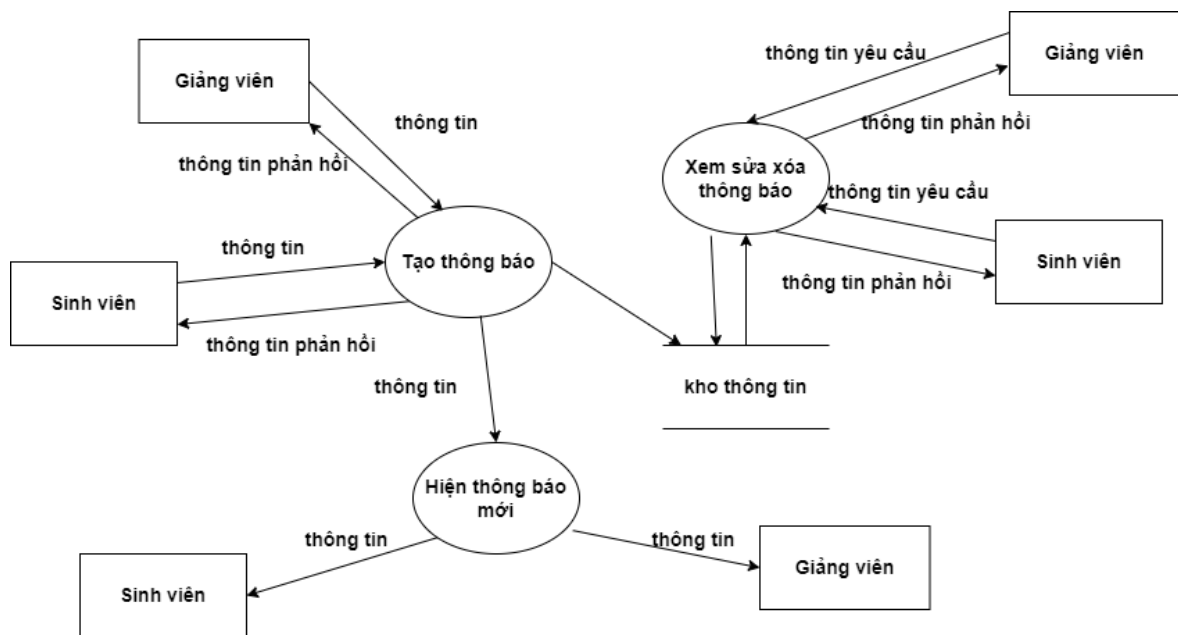
Hình 2. 4 DFD mức dưới đỉnh của quản lý tài khoản.

➤ Quản lý lớp học



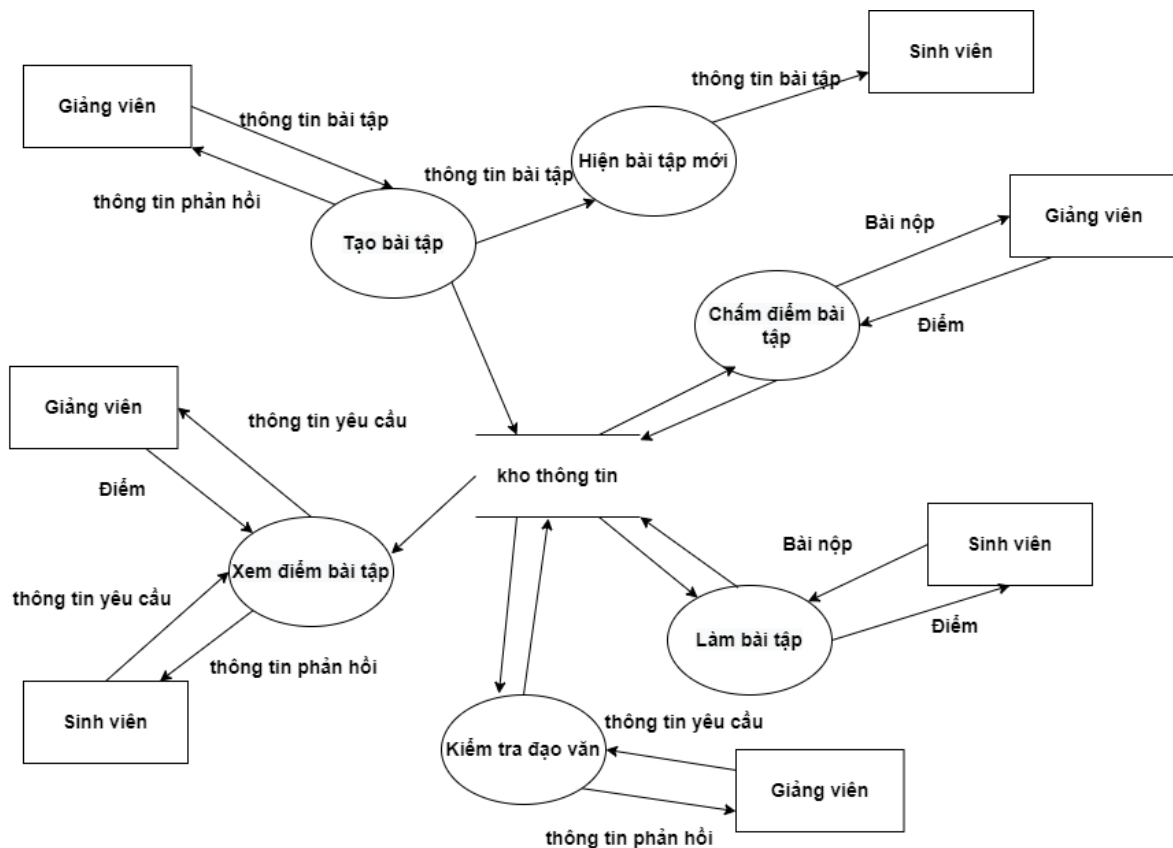
Hình 2. 5 DFD mức dưới đỉnh của quản lý lớp học

➤ Quản lý thông báo



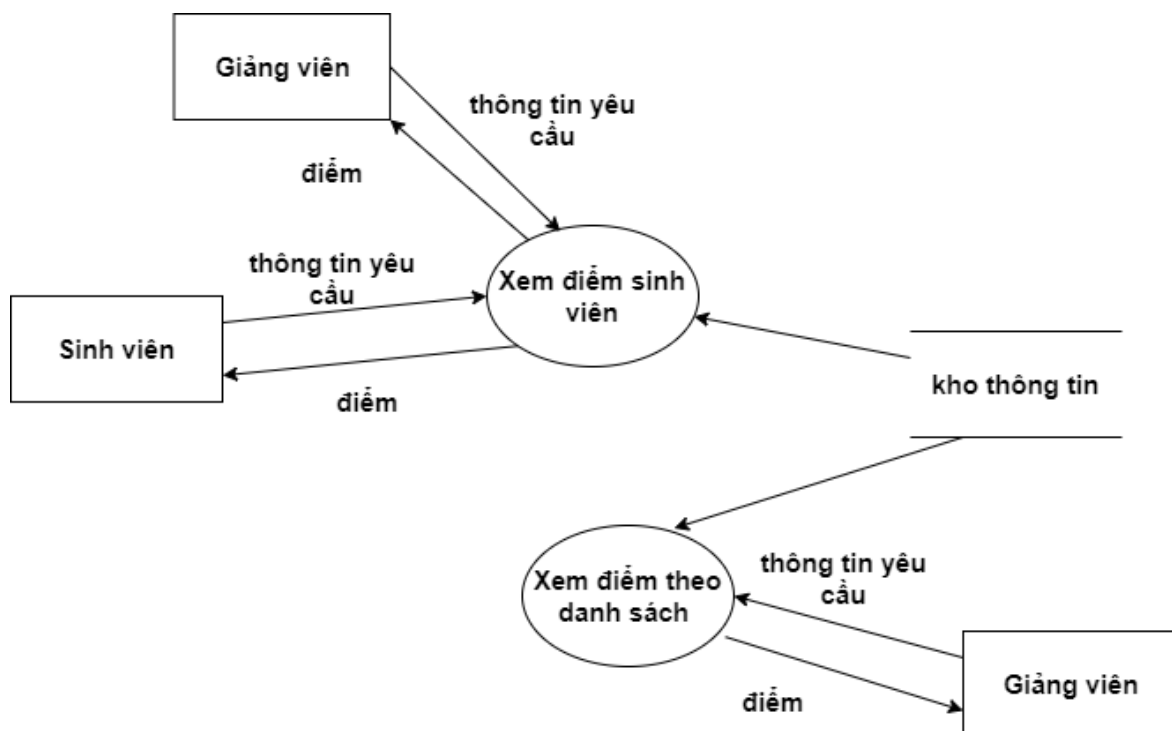
Hình 2. 6 DFD mức dưới đỉnh của quản lý thông báo.

➤ Quản lý bài tập



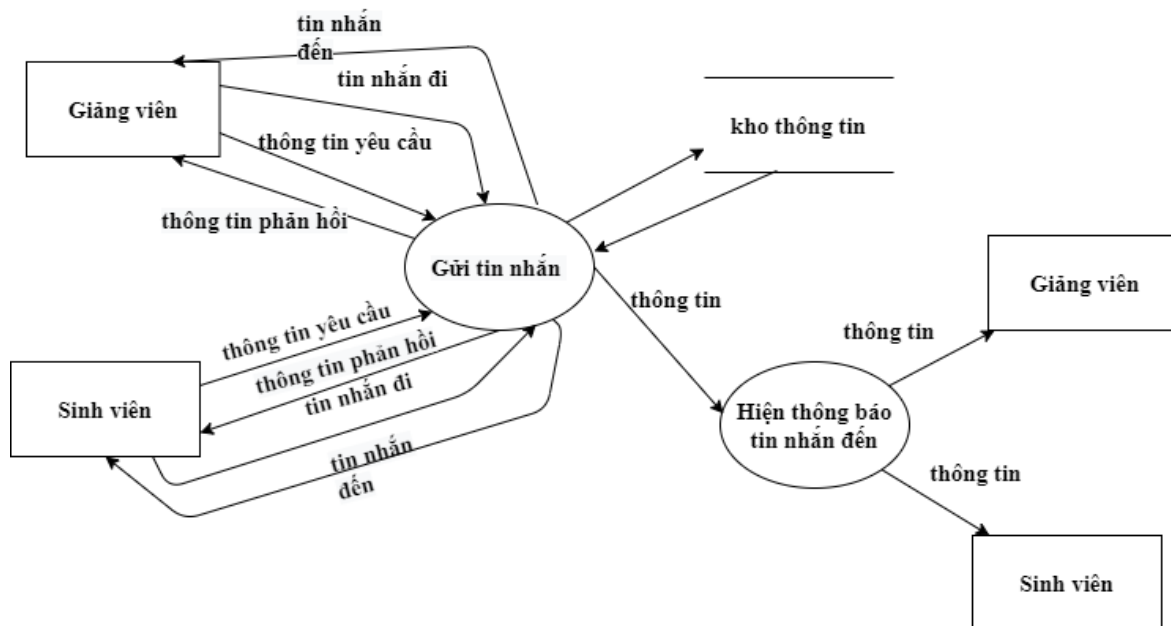
Hình 2. 7 DFD mức dưới đỉnh của quản lý bài tập.

➤ **Thống kê điểm**



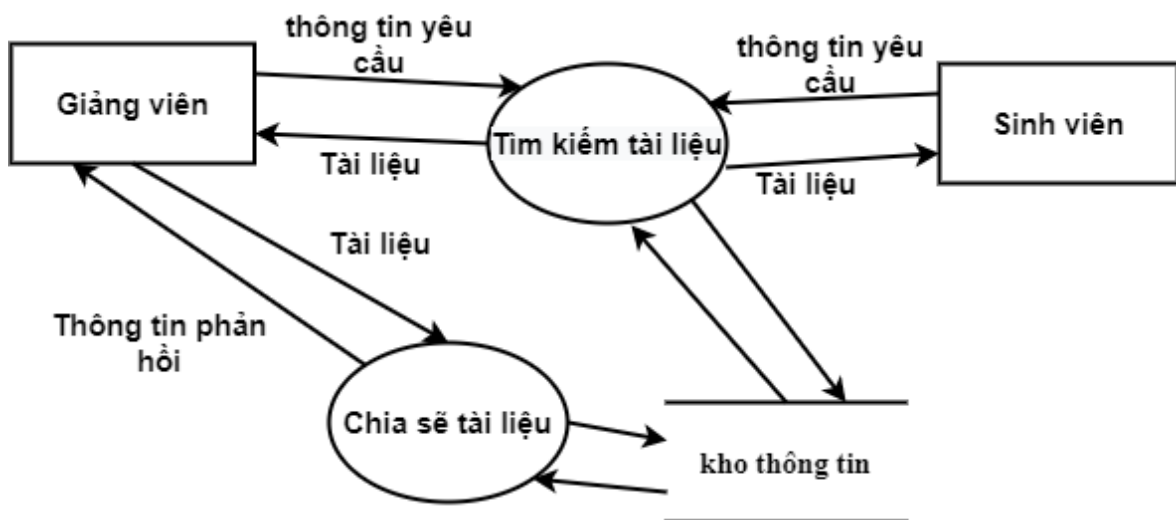
Hình 2. 8 DFD mức dưới đỉnh của thống kê điểm

➤ **Nhắn tin**



Hình 2. 9 DFD mức dưới đỉnh của nhắn tin.

➤ Tài liệu



Hình 2. 10 DFD mức dưới đỉnh của tài liệu.

2.5. Biểu đồ UseCase

2.5.1. Tổng quan hệ thống

Hệ thống LMS cho phép người dùng tổ chức lớp học, các thông tin của hệ thống được đặt ở một sever cố định. Tập trung vào năm quy trình chính:

➤ Quy trình tạo lớp học và tham gia lớp học :

Người dùng chọn chức năng tạo lớp học điền thông tin lớp học cần tạo. Sau khi tạo lớp học thì mỗi lớp học sẽ có một mã lớp học để người tạo lớp có thể chia sẻ để người khác tham gia vào lớp học. Người tạo lớp có thể mời người khác tham gia lớp học bằng email.

➤ Quy trình tạo bài tập và chấm bài tập :

- Đối với bài tập tự luận : Chọn chức năng tạo bài tập và chọn loại bài tập tự luận. Điền thông tin liên quan về bài tập và tạo bài tập. Sau khi sinh viên hoàn tất việc nộp bài tập thì giảng viên chọn bài tập cần chấm điểm, chọn sinh viên cần chấm điểm, xem bài tập của sinh viên và nhập điểm cần chấm.
- Đối với bài tập trắc nghiệm : Giảng viên cần chuẩn bị file câu hỏi trắc nghiệm theo format do chương trình cung cấp. Chọn chức năng tạo bài tập

và chọn loại bài tập trắc nghiệm. Điền thông tin liên quan, chọn file câu hỏi trắc nghiệm và tạo bài tập. Sinh viên sẽ làm bài tập và hệ thống tự chấm điểm và hiện thông tin điểm bài tập.

➤ Quy trình kiểm tra đạo văn :

Sau khi sinh viên đã hoàn tất nộp toàn bộ bài tập tự luận ,giảng viên sẽ chọn bài tập chọn chức năng kiểm tra đạo văn và chờ hệ thống kiểm tra.

➤ Quy trình tải tài liệu và tìm kiếm tài liệu :

Giảng viên chọn tài liệu cần chia sẻ và điền thông tin liên quan tới tài liệu. Muốn tìm kiếm tài liệu thì sinh viên cần điền thông tin muốn tìm và bấm tìm kiếm.

➤ Quy trình tổng kết điểm :

Hệ thống tự tổng kết và đưa ra điểm trung bình của tất cả các bài tập.

➤ Quy trình cấp quyền :

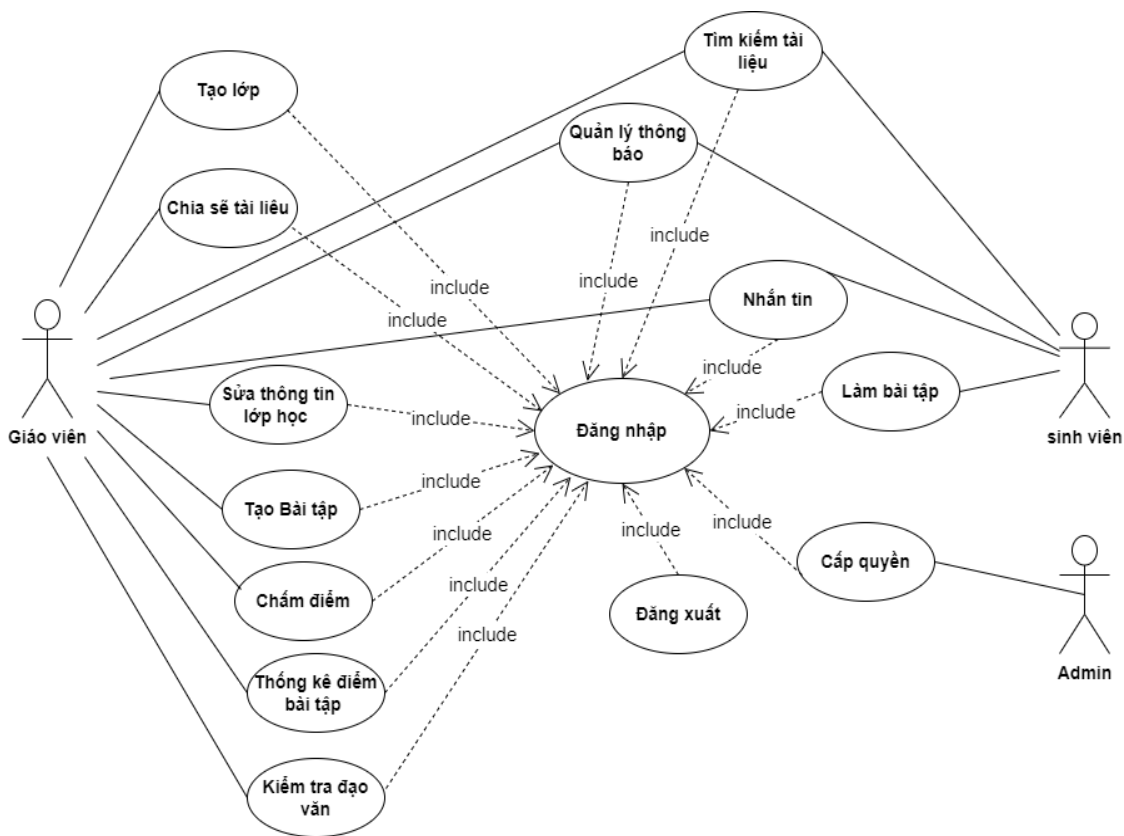
- Đối với quản lý tài khoản : Hệ thống sẽ hiện thông tin danh sách tài khoản admin sẽ thay đổi quyền của tài khoản thành sinh viên hay giáo viên.
- Đối với lớp học : Hệ thống sẽ hiện thông tin danh sách thành viên của lớp học admin sẽ thay đổi quyền của tài khoản đó trong lớp học như thành giáo viên để quản lý lớp học hoặc thành sinh viên.

2.5.2. Các tác nhân

STT	Tên tác nhân	Quyền sử dụng
1	Giảng viên	Tạo bài tập, chấm điểm, chia sẻ tài liệu.
2	Sinh viên	Tạo thông báo, làm bài tập.
2	Admin	Cấp quyền

Bảng 2. 2 Các tác nhân

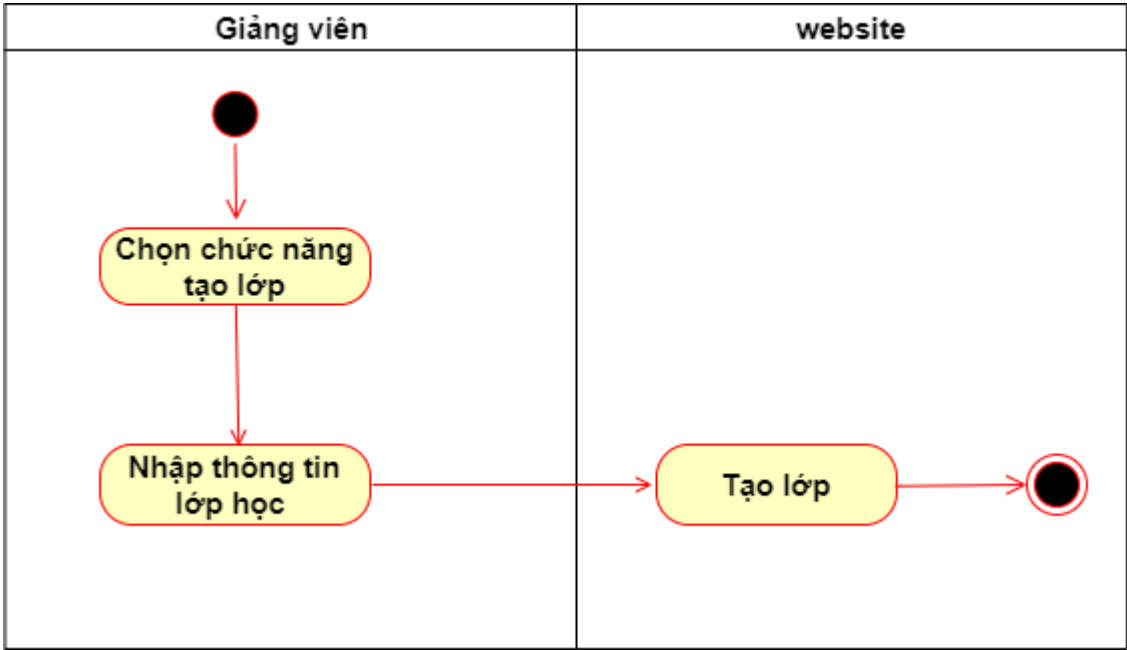
2.5.3. Sơ đồ UseCase tổng quát



Hình 2. 11 Sơ đồ Usecase tổng quát

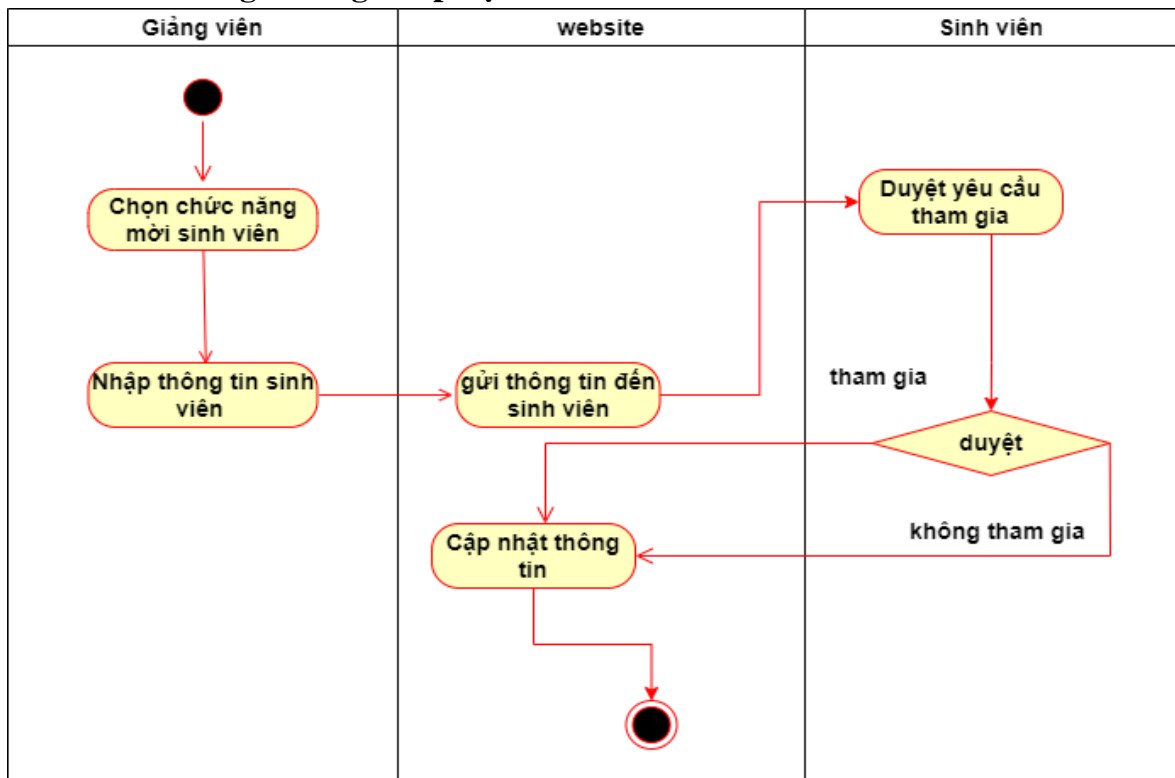
2.6. Biểu đồ hoạt động

2.6.1. Chức năng tạo lớp học



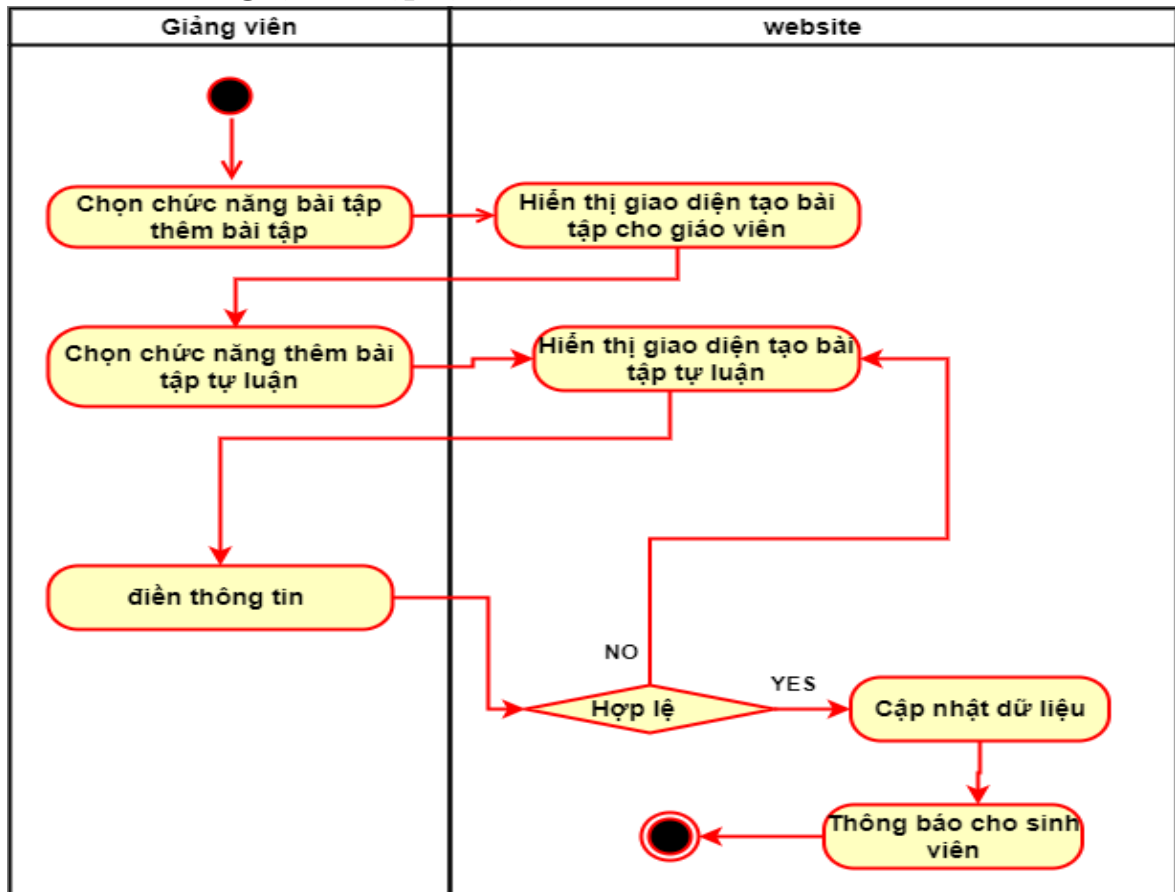
Hình 2. 12 Biểu đồ hoạt động tạo lớp học

2.6.2. Chức năng tham gia lớp học



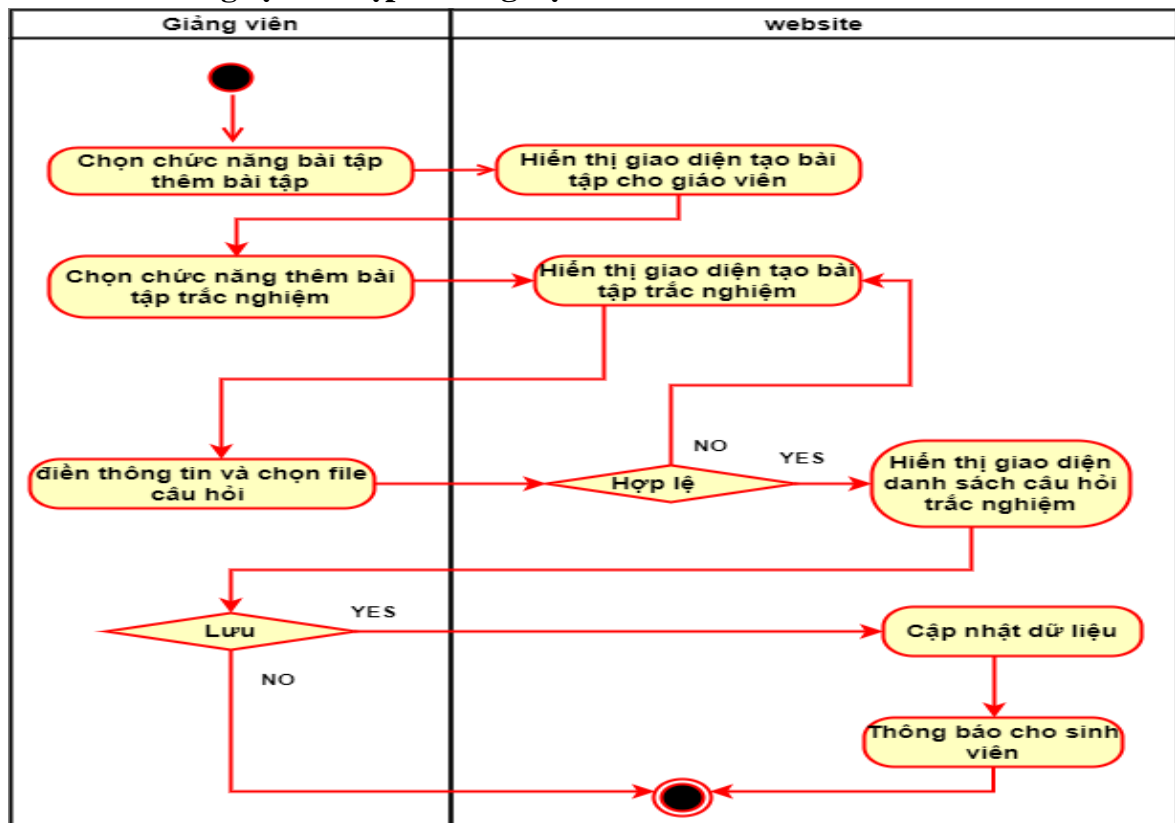
Hình 2. 13 Biểu đồ hoạt động tham gia lớp học

2.6.3. Chức năng tạo bài tập tự luận



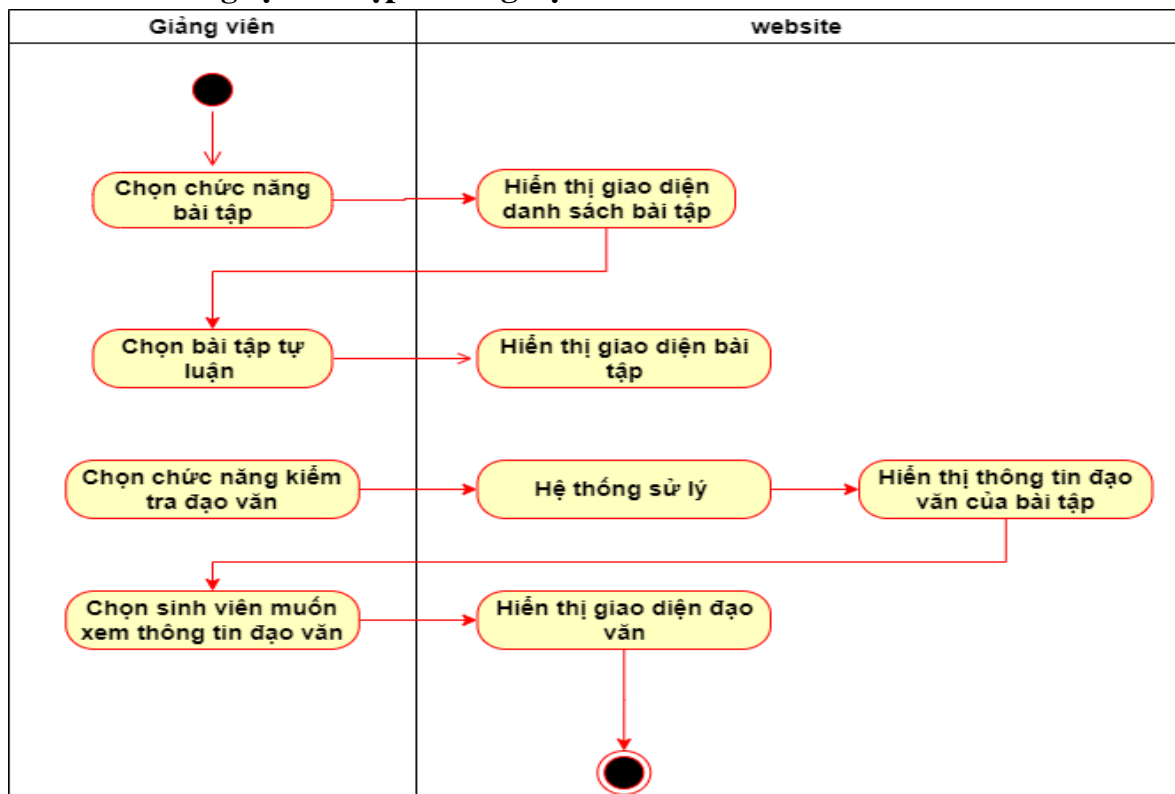
Hình 2. 14 Biểu đồ hoạt động tạo bài tập tự luận

2.6.4. Chức năng tạo bài tập trắc nghiệm



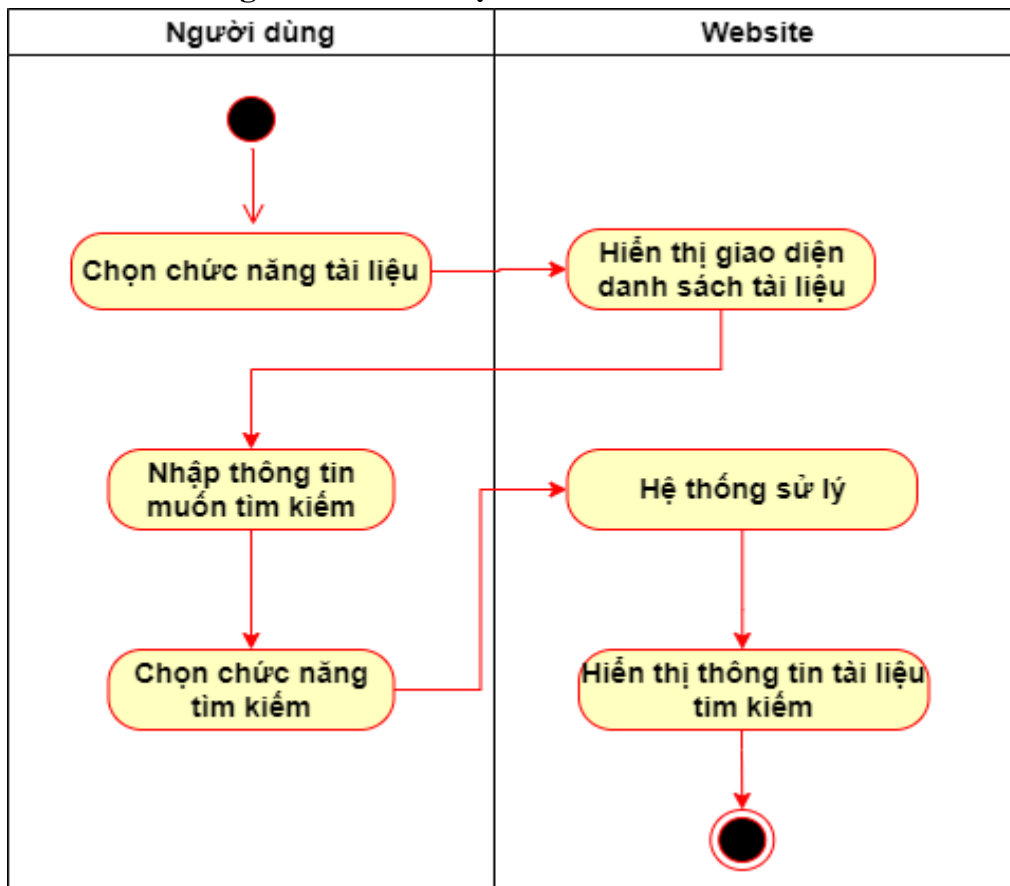
Hình 2. 15 Biểu đồ hoạt động tạo bài tập trắc nghiệm

2.6.5. Chức năng tạo bài tập trắc nghiệm



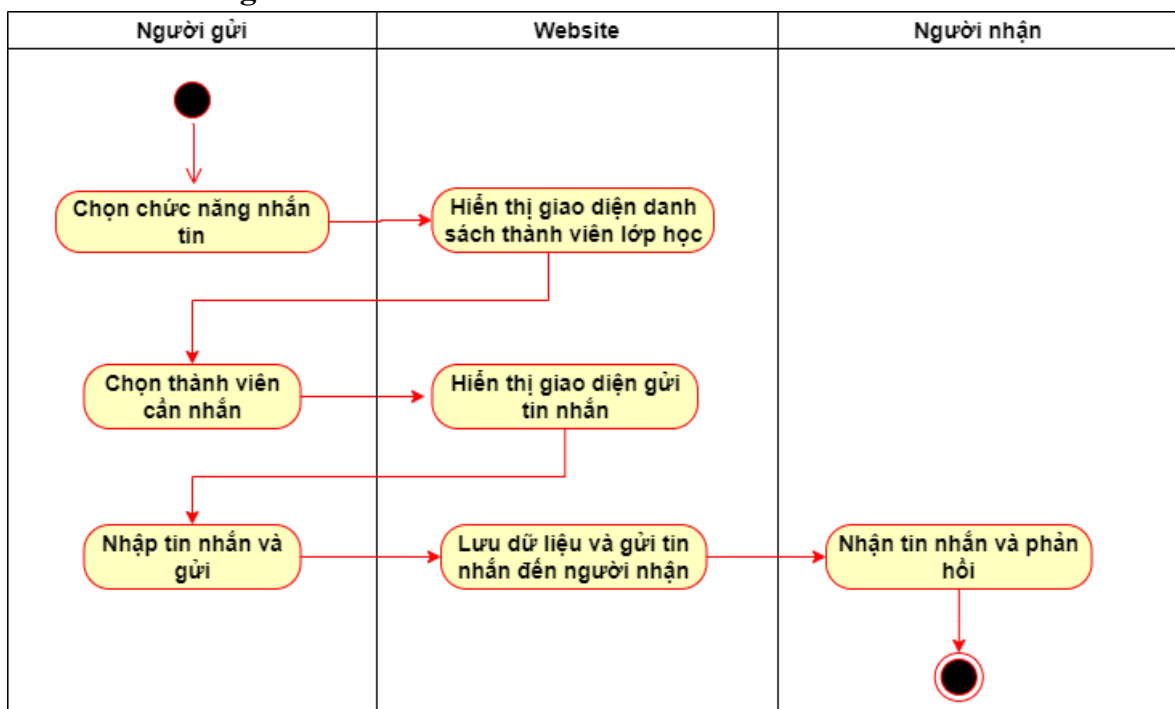
Hình 2. 16 Biểu đồ hoạt động kiểm tra đạo văn

2.6.6. Chức năng tìm kiếm tài liệu



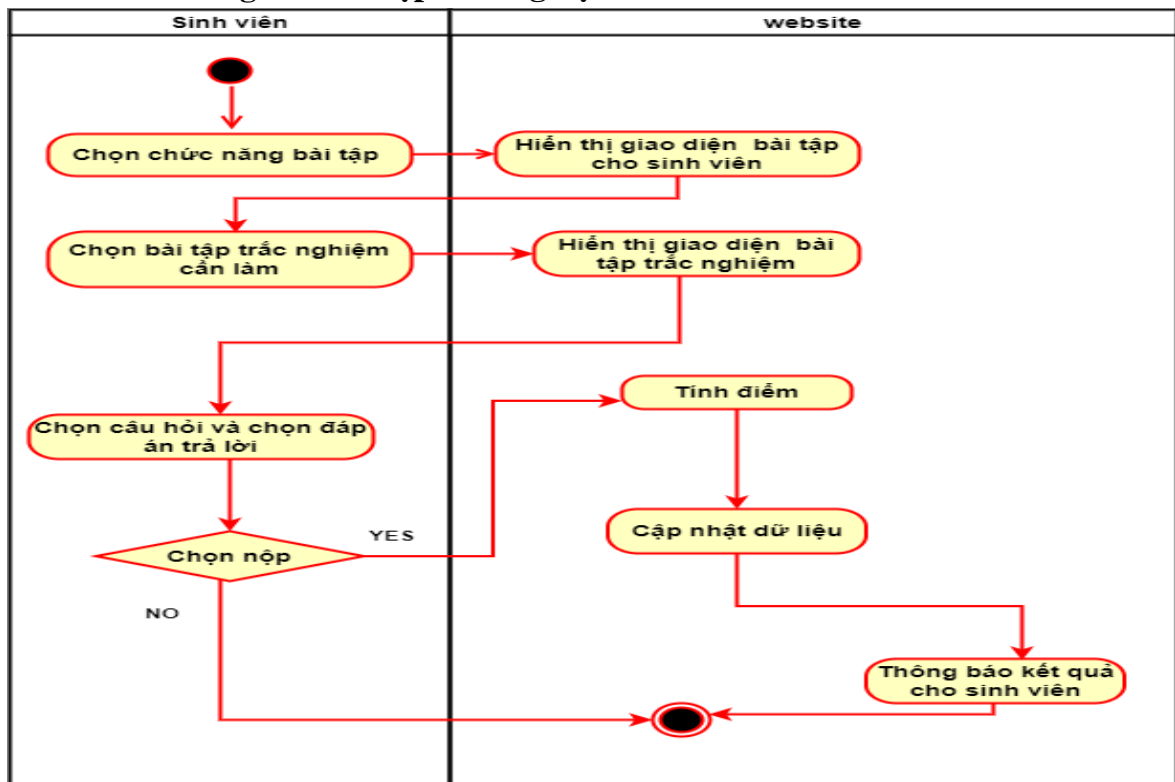
Hình 2. 17 Biểu đồ hoạt động tìm kiếm tài liệu

2.6.7. Chức năng nhắn tin



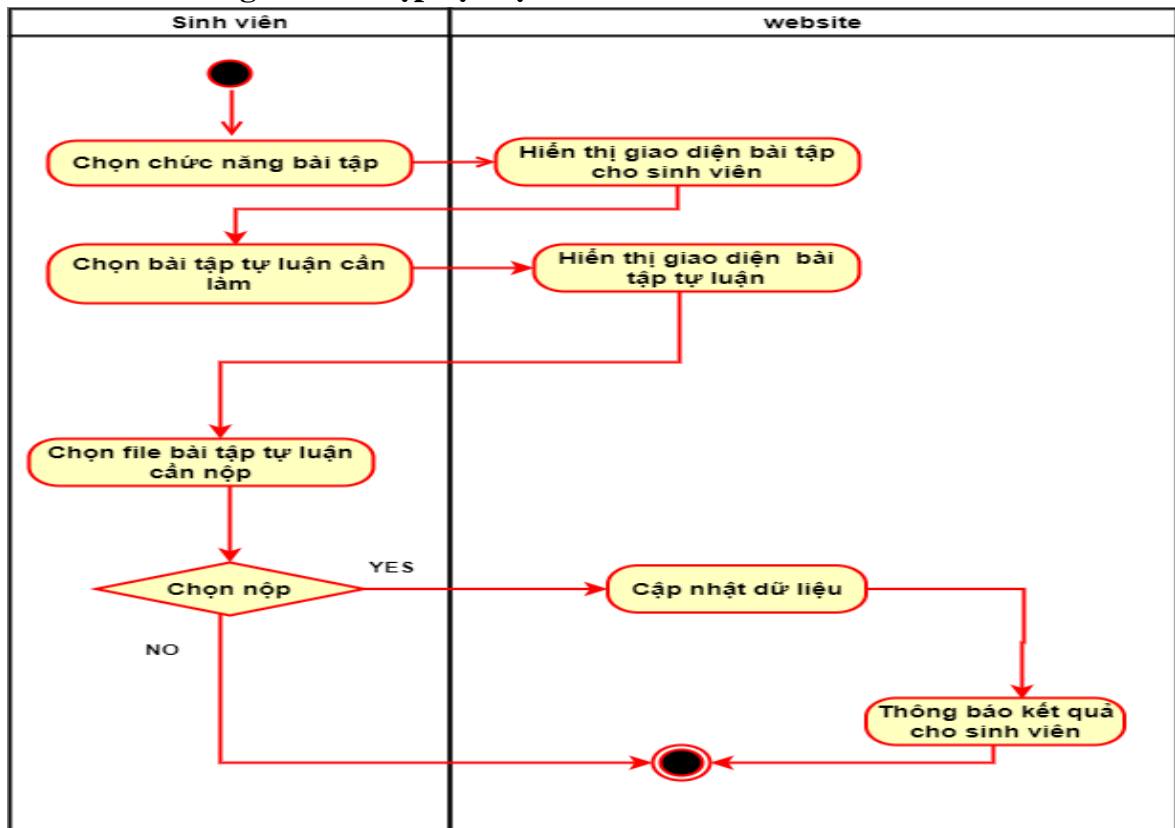
Hình 2. 18 Biểu đồ hoạt động nhắn tin

2.6.8. Chức năng làm bài tập trắc nghiệm



Hình 2. 19 Biểu đồ hoạt động làm bài trắc nghiệm

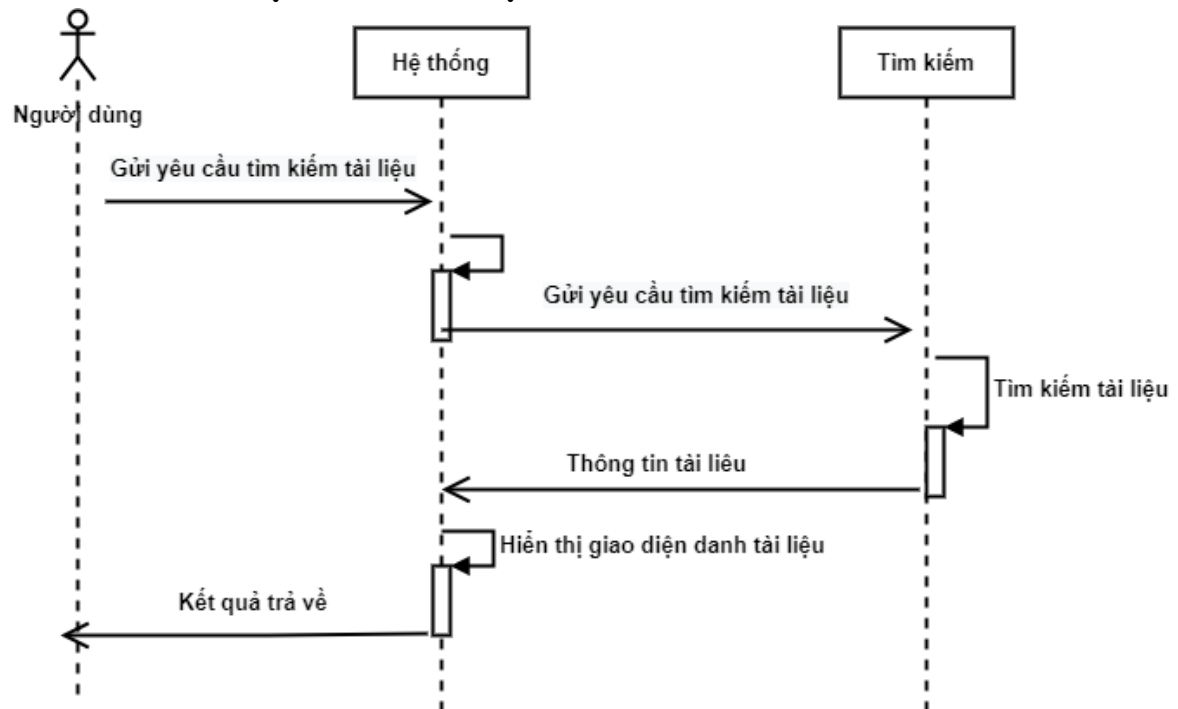
2.6.9. Chức năng làm bài tập tự luận



Hình 2. 20 Biểu đồ hoạt động làm bài tự luận

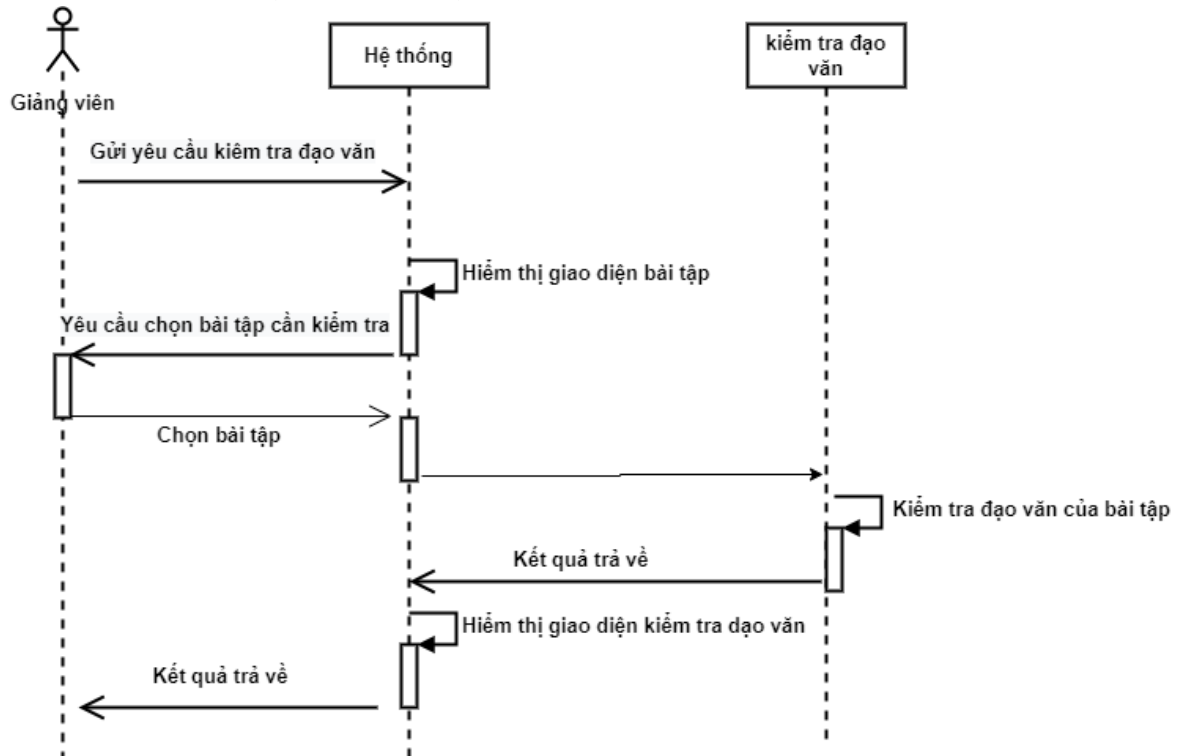
2.7. Sơ đồ tuần tự

2.7.1. Sơ đồ tuần tự tìm kiếm tài liệu



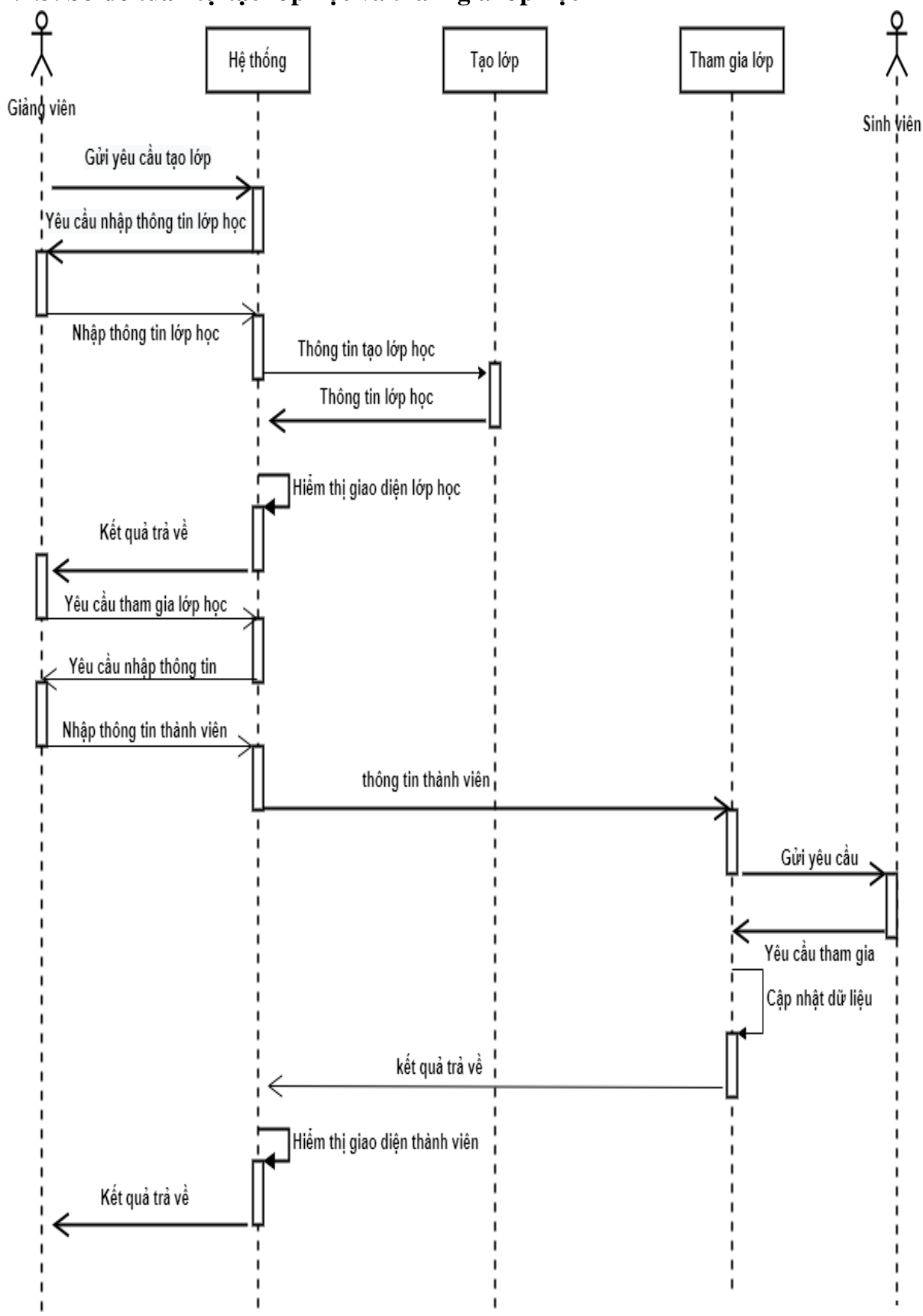
Hình 2. 21 Sơ đồ tuần tự tìm kiếm tài liệu

2.7.2. Sơ đồ tuần tự kiểm tra đạo văn



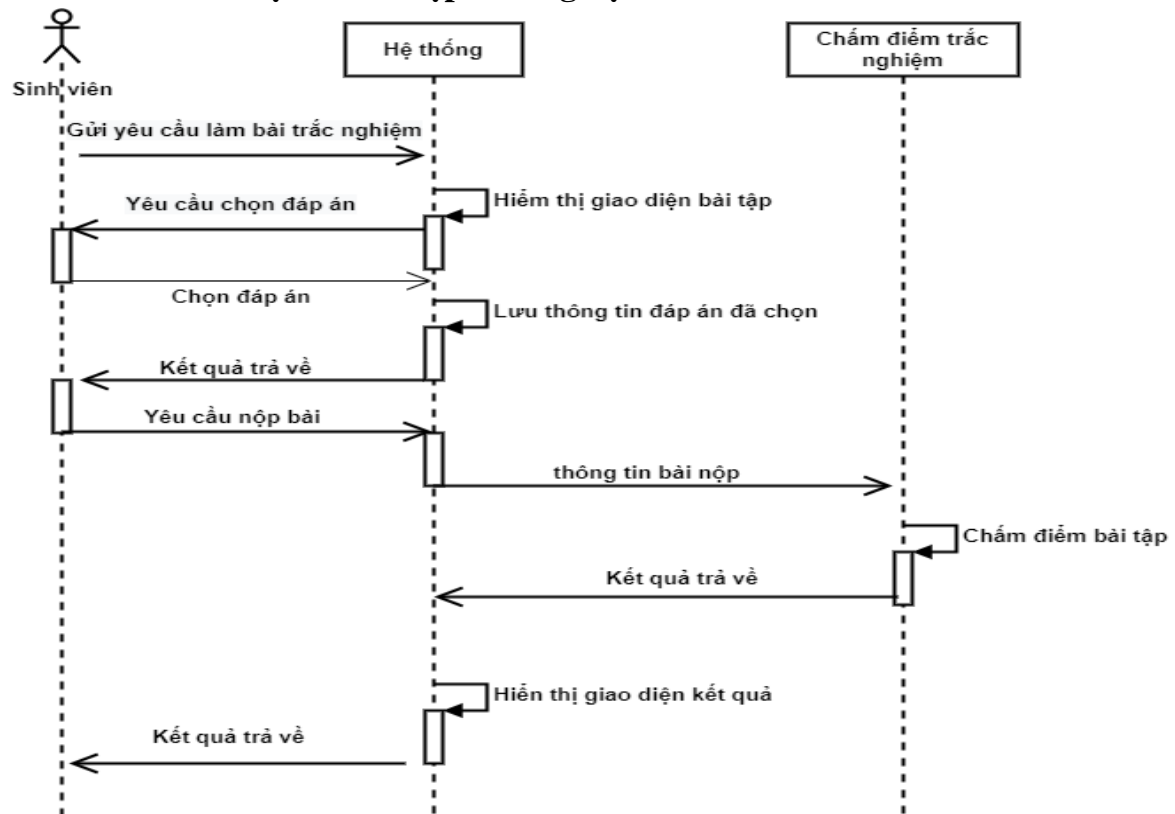
Hình 2. 22 Sơ đồ tuần tự kiểm tra đạo văn

2.7.3. Sơ đồ tuần tự tạo lớp học và tham gia lớp học



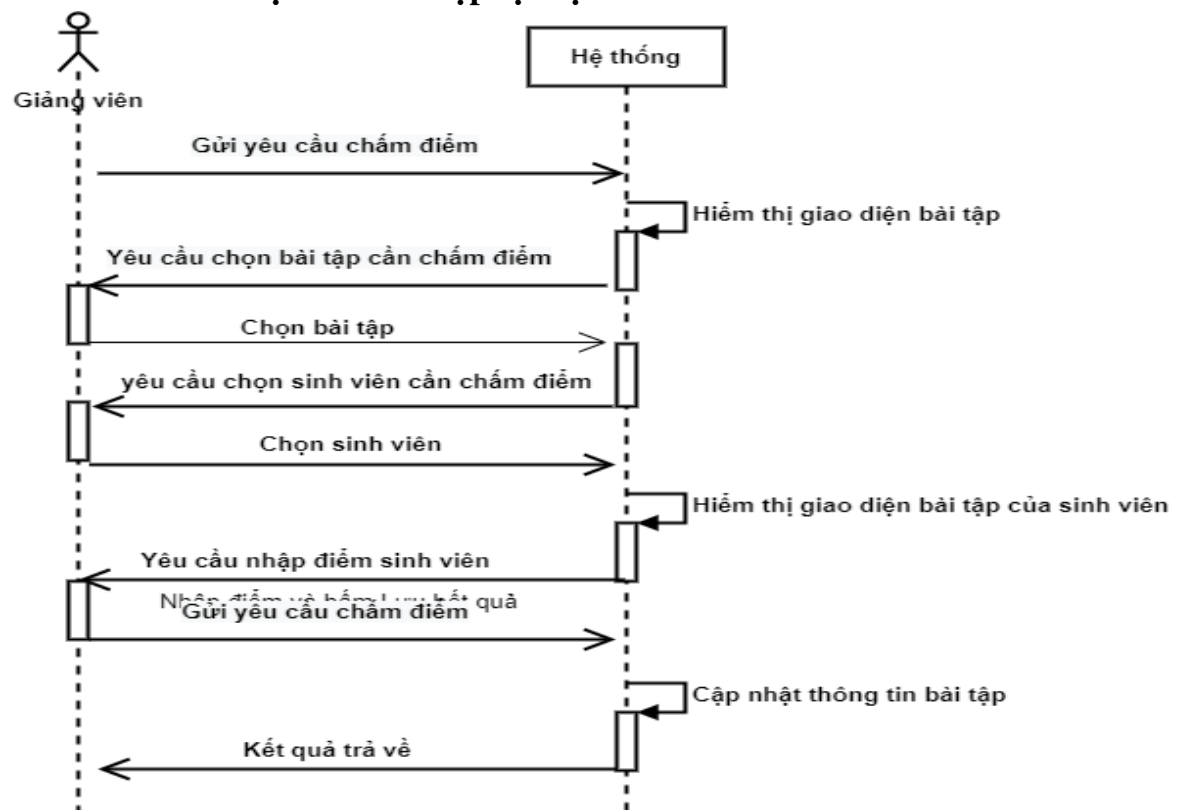
Hình 2. 23 Sơ đồ tuần tự tạo lớp và tham gia lớp học

2.7.4. Sơ đồ tuần tự làm bài tập trắc nghiệm



Hình 2. 24 Sơ đồ tuần tự làm bài trắc nghiệm

2.7.5. Sơ đồ tuần tự chấm bài tập tự luận



Hình 2. 25 Sơ đồ tuần tự chấm bài tập tự luận

CHƯƠNG 3: THIẾT KẾ HỆ THỐNG VÀ TRIỂN KHAI

3.1. Thiết kế hệ thống

3.1.1. Các thực thể và thuộc tính

➤ Tập thực thể tài khoản

Mô Tả : Tài khoản gồm các thuộc tính tên đăng nhập ,mật khẩu ,họ ,tên,email ,hình ảnh,id tài khoản.Mỗi tài khoản được xác định duy nhất bởi tên đăng nhập. Một tài khoản có thể là thành viên của không hoặc nhiều lớp học và có thể tạo không hoặc nhiều lớp học. Một tài khoản không có bài trắc nghiệm ,tự luận nào hoặc có thể có nhiều bài trắc nghiệm tự luận. Một tài khoản có thể không tạo bài tập nào hoặc tạo nhiều bài tập. Một tài khoản có thể không có tin nhắn nào hoặc có nhiều tin nhắn. Một tài khoản có thể không có thông báo nào hoặc có nhiều thông báo.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	tendangnhap	Tên đăng nhập	String
2	matkhau	Mật khẩu	String
4	ho	Họ	String
5	ten	tên	String
6	Email	Email	String
7	hinhanh	Hình ảnh	String
8	Token	Id tài khoản	String

Bảng 3. 1 Thực thể tài khoản

➤ Tập thực thể lớp học

Mô Tả : Lớp học gồm các thuộc tính :Mã lớp học ,tên lớp học ,thông tin lớp học, ngày tạo lớp ,hình ảnh. Mỗi lớp được xác định duy nhất bởi mã lớp học. Một lớp học do một người tạo và có một hoặc nhiều tài khoản là thành viên lớp học tham gia. Một lớp không có thông báo nào hoặc có nhiều thông báo. Một lớp học có thể không có tin nhắn nào hoặc có nhiều tin nhắn thuộc lớp đó. Lớp học có thể không có bài tập nào hoặc có nhiều bài tập.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>malophoc</u>	Mã lớp học	Interger

2	tenlophoc	Tên lớp học	String
3	thongtinlophoc	Thông tin lớp học	String
4	ngaytao	Ngày tạo lớp	Date
5	hinhanh	Hình ảnh	String

Bảng 3. 2 Thực thể lớp học.

➤ **Tập thực thể thông báo**

Mô tả : Một thông báo bao gồm các thông tin : mã thông báo, ngày đăng thông báo ,thông tin thông báo ,loại thông báo. Mỗi thông báo được xác định duy nhất bởi mã thông báo. Một thông báo thì có nhiều file thông báo hoặc không có file thông báo nào. Thông báo do một và chỉ một tài khoản tạo. Một thông báo chỉ thuộc duy nhất một lớp học. Một thông báo có thể là thông báo cho bài tập hoặc là thông báo thông thường.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>mathongbao</u>	Mã thông báo	Interger
2	ngaydang	Ngày đăng	Date
3	thongtin	Thông tin thông báo	String
4	loaithongbao	Loại thông báo	String

Bảng 3. 3 Thực thể thông báo.

➤ **Tập thực thể tin nhắn**

Mô Tả: Tin nhắn gồm các thuộc tính :Mã tin nhắn ,thời gian gửi ,thông tin tin nhắn,người nhận. Mỗi tin nhắn được xác định duy nhất bởi mã tin nhắn. Một tin nhắn thì chỉ thuộc duy nhất 1 lớp. Một tin nhắn do duy nhất 1 tài khoản gửi đi.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>Matinnhan</u>	Mã tin nhắn	Interger
2	nguoinhan	Người nhận	String
3	thongtintinnhan	Thông tin tin nhắn	String
4	thoigiangui	Thời gian gửi	Date

Bảng 3. 4 Thực thể tin nhắn.

➤ **Tập thực thể file thông báo**

Mô tả : file thông báo gồm các thuộc tính : mã file ,tên file ,nơi lưu. Mỗi file thông báo được xác định duy nhất bởi mã file. Một file thông báo thì chỉ thuộc duy nhất một thông báo.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>mafile</u>	Mã file	Interger
2	tenfile	Tên file	String
3	noiluu	Nơi lưu	String

Bảng 3. 5 Thực thể file thông báo.

➤ Tập thực thể bài tập

Mô tả :Bài tập gồm các thuộc tính : Mã bài tập ,chủ đề bài tập ,thông tin bài tập, loại bài tập ,thời gian đăng ,thời gian kết thúc. Mỗi bài tập được xác định duy nhất bởi mã bài tập. Một bài tập thì có không hoặc nhiều file bài tập tự luận. Một bài tập thì có một hoặc nhiều thông báo, thuộc duy nhất một lớp học và do một tài khoản tạo. Một bài tập có không hoặc nhiều bài tập tự luận và bài tập trắc nghiệm. Bài tập thì không có câu hỏi nào hoặc có nhiều câu hỏi.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>mabaitap</u>	Mã bài tập	Interger
2	chude	Chủ đề bài tập	String
3	thongtin	Thông tin bài tập	String
4	loaibaitap	Loại bài tập	String
5	thoigiandang	Thời gian đăng	Date
6	thoigianketthuc	Thời gian kết thúc	Date

Bảng 3. 6 Thực thể bài tập.

➤ Tập thực file bài tập tự luận

Mô tả : file bài tập tự luận gồm các thuộc tính : mã file ,tên file ,nơi lưu. Mỗi file bài tập tự luận được xác định duy nhất bởi mã file. Một file bài tập tự luận thì chỉ thuộc duy nhất một bài tập.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>mafile</u>	Mã file	Interger

2	tenfile	Tên file	String
3	noiluu	Nơi lưu	String

Bảng 3. 7 Thực thể file bài tập tự luận.

➤ **Tập thực thể thông tin bài tập tự luận**

Mô tả : thông tin bài tập tự luận gồm các thuộc tính : mã file ,tên file ,nơi lưu, ngày nộp, kiểm tra đạo văn, dữ liệu file. Mỗi thông tin bài tập tự luận được xác định duy nhất bởi mã file. Một thông tin bài tập tự luận thì chỉ thuộc duy nhất một bài tập tự luận. Một thông tin bài tập tự luận có nhiều kiểm tra đạo văn.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>mafile</u>	Mã file	Integer
2	tenfile	Tên file	String
3	noiluu	Nơi lưu	String
4	ngaynop	Ngày nộp	Date
5	kiemtradaovan	Kiểm tra đạo văn	Boolean
6	dulieufile	Dữ liệu file	String

Bảng 3. 8 Thực thể thông tin bài tập tự luận.

➤ **Tập thực thể đáp án**

Mô tả : đáp án gồm các thuộc tính : mã đáp án ,nội dung ,loại đáp án. Mỗi đáp án được xác định duy nhất bởi mã đáp án. Một đáp án thì thuộc nhiều bài tập trắc nghiệm và thuộc duy nhất một câu hỏi.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>madapan</u>	Mã đáp án	Integer
2	noidung	Nội dung	String
3	loaidapan	Loại đáp án	String

Bảng 3. 9 Thực thể đáp án.

➤ **Tập thực thể câu hỏi**

Mô tả : Câu hỏi gồm các thuộc tính : mã câu hỏi ,nội dung ,ngày thêm. Mỗi câu hỏi được xác định duy nhất bởi mã câu hỏi. Một câu hỏi có một hoặc nhiều đáp án

trong một câu hỏi. Một câu hỏi thì thuộc duy nhất trong một bài tập, một câu hỏi thuộc nhiều bài tập trắc nghiệm.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>macauhoi</u>	Mã câu hỏi	Interger
2	noidung	Nội dung	String
3	ngaythem	Ngày thêm	Date

Bảng 3. 10 Thực thể câu hỏi.

➤ **Tập thực thể bài tập trắc nghiệm**

Mô tả : Một bài tập trắc nghiệm gồm các thuộc tính : mã bài nộp ,ngày nộp ,trạng thái nộp bài ,điểm bài tập. Mỗi bài tập trắc nghiệm được xác định duy nhất bởi mã bài nộp. Một bài tập trắc nghiệm thì thuộc một bài tập và có duy nhất một tài khoản nộp. Một bài tập có một hoặc nhiều câu hỏi và đáp án.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>mabainop</u>	Mã bài nộp	Interger
2	ngaynop	Ngày nộp bài tập	Date
3	trangthai	Trạng thái nộp bài	Boolean
4	diem	Điểm bài tập	Float

Bảng 3. 11 Thực thể bài tập trắc nghiệm.

➤ **Tập thực thể bài tập tự luận**

Mô tả : Một bài tập tự luận gồm các thuộc tính : mã bài nộp, ngày nộp, trạng thái nộp bài ,điểm bài tập. Mỗi bài tập tự luận được xác định duy nhất bởi mã bài nộp. Một bài tập tự luận thì thuộc duy nhất một bài tập và có duy nhất một tài khoản nộp. Một bài tập tự luận có một hoặc nhiều thông tin bài tập tự luận.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>mabainop</u>	Mã bài nộp	Interger
2	ngaynop	Ngày nộp bài tập	Date
3	trangthai	Trạng thái nộp bài	Boolean
4	diem	Điểm bài tập	Float

Bảng 3. 12 Thực thể bài tập tự luận.

➤ **Tập thực thể bình luận bài tập tự luận**

Mô tả : Một bình luận bài tập tự luận gồm các thuộc tính : nội dung ,thời gian bình luận. Mỗi bình luận bài tập tự luận được xác định duy nhất bởi mã bình luận. Một bình luận bài tập tự luận thì thuộc duy nhất một bài tập tự luận và có duy nhất một tài khoản. Một bình luận bài tập tự luận có một hoặc nhiều trả lời bình luận bài tập tự luận.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>ma</u>	Mã bình luận	Interger
2	noidung	Nội dung	String
3	thoigianbinhluan	thời gian bình luận	Date

Bảng 3. 13 Thực thể bình luận bài tập trắc nghiệm.

➤ **Tập thực thể bình luận bài tập trắc nghiệm**

Mô tả : Một bình luận bài tập trắc nghiệm gồm các thuộc tính : nội dung ,thời gian bình luận. Mỗi bình luận bài tập trắc nghiệm được xác định duy nhất bởi mã bình luận. Một bình luận bài tập trắc nghiệm thì thuộc duy nhất một bài tập trắc nghiệm và có duy nhất một tài khoản. Một bình luận bài tập trắc nghiệm có một hoặc nhiều trả lời bình luận bài tập trắc nghiệm.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>ma</u>	Mã bình luận	Interger
2	noidung	Nội dung	String
3	thoigianbinhluan	thời gian bình luận	Date

Bảng 3. 14 Thực thể bình luận bài tập trắc nghiệm.

➤ **Tập thực thể bình luận thông báo**

Mô tả : Một bình luận thông báo gồm các thuộc tính : nội dung ,thời gian bình luận. Mỗi bình luận thông báo được xác định duy nhất bởi mã bình luận. Một bình luận thông báo thì thuộc duy nhất một thông báo và có duy nhất một tài khoản. Một bình luận thông báo có một hoặc nhiều trả lời bình luận thông báo.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
-----	------------	-------	--------------

1	<u>ma</u>	Mã bình luận	Interger
2	noidung	Nội dung	String
3	thoigianbinhluan	thời gian bình luận	Date

Bảng 3. 15 Thực thể bình luận thông báo.

➤ **Tập thực thể trả lời bình luận bài tập tự luận**

Mô tả : Một trả lời bình luận bài tập tự luận gồm các thuộc tính : nội dung ,thời gian bình luận. Mỗi trả lời bình luận bài tập tự luận được xác định duy nhất bởi mã bình luận. Một trả lời bình luận bài tập tự luận thì thuộc duy nhất một bình luận bài tập tự luận và có duy nhất một tài khoản.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>ma</u>	Mã bình luận	Interger
2	noidung	Nội dung	String
3	thoigianbinhluan	thời gian bình luận	Date

Bảng 3. 16 Thực thể bình luận bài tập tự luận.

➤ **Tập thực thể bình trả lời luận bài tập trắc nghiệm**

Mô tả : Một trả lời bình luận bài tập trắc nghiệm gồm các thuộc tính : nội dung ,thời gian bình luận. Mỗi trả lời bình luận bài tập trắc nghiệm được xác định duy nhất bởi mã bình luận. Một trả lời bình luận bài tập trắc nghiệm thì thuộc duy nhất một bình luận bài tập trắc nghiệm và có duy nhất một tài khoản.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>ma</u>	Mã bình luận	Interger
2	noidung	Nội dung	String
3	thoigianbinhluan	thời gian bình luận	Date

Bảng 3. 17 Thực thể bình luận bài tập trắc nghiệm.

➤ **Tập thực thể trả lời bình luận thông báo**

Mô tả : Một trả lời bình luận thông báo gồm các thuộc tính : nội dung ,thời gian bình luận. Mỗi trả lời bình luận thông báo được xác định duy nhất bởi mã bình luận. Một trả lời bình luận thông báo thì thuộc duy nhất một bình luận thông báo và có duy nhất một tài khoản.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>ma</u>	Mã bình luận	Interger
2	noidung	Nội dung	String
3	thoigianbinhluan	thời gian bình luận	Date

Bảng 3. 18 Thực thể trả lời bình luận thông báo.

➤ **Tập thực thể kiểm tra đạo văn**

Mô tả : Một kiểm tra đạo văn gồm thuộc tính : phần trăm. Mỗi kiểm tra đạo văn được xác định duy nhất bởi mã kiểm tra. Một kiểm tra đạo văn thì thuộc duy nhất một thông tin bài tập tự luận.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>ma</u>	Mã kiểm tra	Interger
2	phantram	Phần trăm	Float

Bảng 3. 19 Thực thể kiểm tra đạo văn.

➤ **Tập thực thể tài liệu**

Mô tả : Một tài liệu gồm thuộc tính : tên, vị trí lưu,nội dung,hình ảnh ,lượt tải xuống ,ngày tải lên ,lượt xem. Mỗi tài liệu được xác định duy nhất bởi mã tài liệu. Một tài liệu thì thuộc duy nhất một lớp học và do một tài khoản tải lên.

STT	Thuộc tính	Mô tả	Kiểu dữ liệu
1	<u>ma</u>	Mã tài liệu	Interger
2	ten	Tên	String
3	vitriluu	Vị trí lưu	String
4	noidung	Nội dung	String
5	hinhanh	Hình ảnh	String
6	luottai	Lượt tải xuống	Interger
7	luotxem	Luot xem	Interger
8	ngaytailen	Ngày tải lên	Date

Bảng 3. 20 Thực thể tài liệu.

3.1.2. Xây dựng mô hình thực thể liên kết

3.3.3. Mô hình thực thể liên kết sang mô hình quan hệ

Từ mô hình thực thể liên kết sau khi chuyển đổi ta có các bảng sau :

- Taikhoan(**tendangnhap**, matkhou, ho, ten, email, hinhanh, token)
 - tendangnhap là khóa chính và dùng để lưu thông tin tài khoản.
 - matkhou, ho, ten, email, hinhanh, token: lưu mật khẩu, họ, tên, địa chỉ email, đường dẫn hình ảnh, và token của tài khoản
- Lophoc(**malop**, tenlop, thongtin, ngaytao, hinhanh, tendangnhap)
 - malop là khóa chính của bảng Lophoc
 - tenlop, thongtin, ngaytao, hinhanh : các thuộc tính của bảng Lophoc. Dùng để lưu thông tin như tên lớp ,thông tin chi tiết, ngày tạo lớp học, hình ảnh lớp học.
 - tendangnhap là khóa ngoài của bảng Lophoc dung để liên kết với bảng Taikhoan, được dùng để xác định người tạo lớp học.
- Thanhvienlop(**malop**, **tendangnhap**, ngaythamgia)
 - malop và tendangnhap là khoa chính của bảng thành viên lớp, malop dùng để xác nhận thành viên thuộc lớp học nào, tendangnhap dùng để xác nhận thành viên đó là thành viên nào.
 - ngaythamgia dùng để xác nhận ngày tham gia lớp học của thành viên đó.
- Tinnhan(**matinnhan**, **malop**, **tendangnhap**, nguoinhan, thoigianguoi,thongtin)
 - matinnhan là khóa chính của bảng Tinnhan
 - malop và tendangnhap là khóa ngoại của bảng Tinnhan, malop xác nhận tin nhắn thuộc lớp nào, tendangnhap xác nhận ai là người gửi tin.
 - nguoinhan, thoigianguoi ,thongtin: thông tin người nhận tin, thời gian gửi tin và nội dung tin nhắn.
- Filethongbao(**mafile**, tenfile, noiluu, **mabaithongbao**)
 - mafile: khóa chính của bảng
 - mabaithongbao là khóa ngoài của bảng Filethongbao, xác định file này thuộc thông báo nào.
 - tenfile và nơi lưu :tên file thông báo và vị trí lưu trữ file.

- Thongbao(**mabaithongbao**, ngaydang, thongtin, loaithongbao, **malop**, **mabaitap**, **tendangnhap**)
 - mabaithongbao :_khóa chính của bảng
 - ngaydang, thongtin, loaithongbao : ngày đăng thông báo, nội dung thông báo và loại thông báo. Có hai loại thông báo : thông báo lớp học và thông báo bài tập.
 - malop, mabaitap và tendangnhap : malop xác định thông báo thuộc lớp nào, mabaitap sẽ xác nhận thông báo của bài tập nào, tendangnhap xác định ai là người tạo thông báo.
- Filebaitaptuluan(**mafile**, tenfile, noiluu, **mabaitap**)
 - mafile : khóa chính của bảng
 - mabaitap: xác định file này là của bài tập nào.
 - tenfile và noiluu : tên file bài tập tự luận và vị trí lưu trữ file.
- Baitap(**mabaitap**, chude, thongtin, loaibaitap, thoigiandang, thoigianketthuc, **malop**, **tendangnhap**)
 - mabaitap: khóa chính của bảng
 - chude, thongtin, loaibaitap, thoigiandang, thoigianketthuc dùng để xác định chủ đề bài tập, nội dung bài tập, loại bài tập là trắc nghiệm hay tự luận, thời gian tạo bài tập và thời gian hoàn thành của bài tập.
 - malop và tendangnhap là khóa ngoại của bảng và để xác định bài tập của lớp nào và ai là người tạo bài tập.
- Thongtinbaitaptuluan(**mafile**, tenfile, noiluu, ngaynop, kiemtradaovan, **dulieufile**, **mabainop**)
 - mafile: khóa chính của bảng
 - tenfile, noiluu, ngaynop, kiemtradaovan, **dulieufile** : dùng để xác định thông tin file nộp bài tập như tên file, vị trí lưu file, ngày nộp bài, có kiểm tra đạo văn hay chưa và nội dung của file.
 - mabainop : là khóa ngoại để xác định thuộc bài tập nào và sinh viên đã nộp bao nhiêu file.
- Baitaptuluan(**mabainop**, ngaynop, trangthai, diem, **tendangnhap**, **mabaitap**)
 - mabainop: là khóa chính của bảng Baitaptuluan

- ngaynop, trangthai, diem để xác định được đã nộp bài vào ngày nào, trạng thái là đã nộp hay chưa và điểm của bài tập.
 - tendangnhap và mabaitap là khóa ngoại của bảng để xác định ai là người nộp bài và bài đó thuộc bài tập nào
- Baitaptracnghiem(**mabainop**, ngaynop, trangthai, diem, **tendangnhap**, **mabaitap**)
- mabainop: là khóa chính của bảng baitaptracnghiem
 - ngaynop, trangthai và diem để xác định ngày hoàn thành bài tập, trạng thái của bài tập là đã nộp hay chưa và điểm bài tập.
 - tendangnhap và mabaitap là khóa ngoại của bảng để xác định ai là người nộp bài và bài đó thuộc bài tập nào
- Cauhoi(**macauhoi**,noidung,ngaythem,**mabaitap**)
- macauhoi: là khóa chính của bảng Cauhoi
 - mabaitap: là khóa ngoại của bảng, được dùng để xác định câu hỏi này là thuộc bài tập trắc nghiệm nào.
 - noidung và ngaythem để lưu trữ nội dung câu hỏi và ngày thêm câu hỏi.
- Dapan(**madapan**,noidung,loaidapan,**macauhoi**)
- madapan: là khóa ngoại của bảng Dapan.
 - macauhoi: là khóa ngoại để xác định đáp án này là của câu hỏi nào.
 - noidung và loaidapan để lưu nội dung của đáp án trắc nghiệm và loại đáp án là đáp án đúng hay sai.
- Thongtinbaitaptracnghiem(**mabainop**,**macauhoi**,**madapan**)
- mabainop,macauhoi,madapan: là khóa chính của của thông tin bài tập, mabainop để xác định câu hỏi trắc nghiệm này là ai làm, macauhoi để xác định đây là câu hỏi nào và madapan xác định đáp án đã chọn
- Binhluanbaitaptuluan(**ma**, noidung, thoigianbinhluan ,**mabainop**,**tendangnhap**)
- ma: khóa chính của bảng Binhluanbaitaptuluan
 - mabainop và tendangnhap là khóa ngoại của bảng để xác định bình luận thuộc bài tập nào và ai là người tạo bình luận đó.
 - noidung và thoigianbinhluan dùng để lưu trữ thông tin nội dung bình luận và thời gian tạo.

- Binhluanbaitaptracnghiem (ma, noidung, thoigianbinhluan ,mabainop,tendangnhap)
 - ma: khóa chính của bảng Binhluanbaitaptuluan
 - mabainop và tendangnhap là khóa ngoại của bảng để xác định bình luận thuộc bài tập nào và ai là người tạo bình luận đó.
 - noidung và thoigianbinhluan dùng để lưu trữ thông tin nội dung bình luận và thời gian tạo.
- Binhluanthongbao (ma, noidung, thoigianbinhluan ,mathongbao,tendangnhap)
 - ma: khóa chính của bảng Binhluanbaitaptuluan
 - mathongbao và tendangnhap là khóa ngoại của bảng để xác định bình luận thuộc thông báo nào và ai là người tạo bình luận đó.
 - noidung và thoigianbinhluan dùng để lưu trữ thông tin nội dung bình luận và thời gian tạo.
- Traloibinhluanbaitaptuluan(ma, noidung, thoigianbinhluan ,mabinhluanbaitaptuluan,tendangnhap)
 - ma: khóa chính của bảng Binhluanbaitaptuluan
 - mabinhluanbaitaptuluan và tendangnhap là khóa ngoại của bảng để xác định đã trả lời cho bình luận thuộc bài tập nào và ai là người tạo bình luận đó.
 - noidung và thoigianbinhluan dùng để lưu trữ thông tin nội dung bình luận và thời gian tạo.
- Traloibinhluanbaitaptracnghiem (ma, noidung, thoigianbinhluan ,mabinhluanbaitaptracnghiem,tendangnhap)
 - ma: khóa chính của bảng Binhluanbaitaptuluan
 - mabinhluanbaitaptracnghiem và tendangnhap là khóa ngoại của bảng để xác định đã trả lời cho bình luận thuộc bài tập nào và ai là người tạo bình luận đó.
 - noidung và thoigianbinhluan dùng để lưu trữ thông tin nội dung bình luận và thời gian tạo.
- Traloibinhluanthongbao (ma, noidung, thoigianbinhluan ,mabinhluanthongbao,tendangnhap)

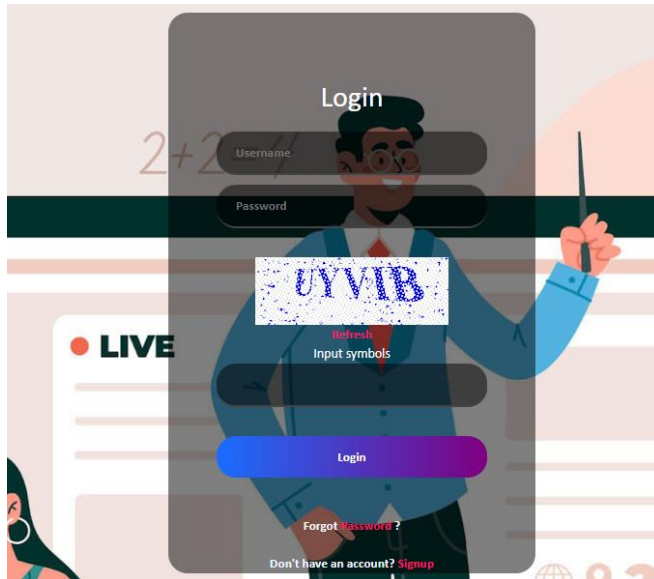
- ma: khóa chính của bảng Binhluanbaitaptuluan
 - mabinhluanthongbao và tendangnhap là khóa ngoại của bảng để xác định đã trả lời cho bình luận thuộc thông báo nào và ai là người tạo bình luận đó.
 - noidung và thoigianbinhluan dùng để lưu trữ thông tin nội dung bình luận và thời gian tạo.
- kiemtradaovan(ma, phantram, mafile)
- ma: là khóa ngoại của bảng kiemtradaovan
 - phantram: lưu thông tin phần tram đạo văn của file
 - mafile : là khóa ngoại của bảng xác định đã kiểm tra đạo văn cho file nào
- tailieu(ma ,ten, vitriluu ,noidung ,hinhanh ,luottai ,luotxem, ngaytailen, malop, tendangnhap)
- ma: là khóa chính bảng tailieu
 - ten, vitriluu, noidung, hinhanh, luottai, luotxem, ngaytailen dùng để lưu thông tin như tên file ,nơi lưu file, nội dung file, hình ảnh giới thiệu của file, lượt truy cập file, lượt tải file, ngày thêm file
 - malop và tendangnhap là khóa ngoại của bảng để xác định file thuộc lớp nào và ai đã tải file lên hệ thống.

3.2 Triển khai

Sau đây là một vài giao diện hệ thống chủ yếu:

3.2.1 Giao diện đăng nhập

Người sử dụng cần đăng nhập để sử dụng các chức năng của hệ thống. Chức năng đăng nhập yêu cầu tên đăng nhập, mật khẩu và mã xác nhận.

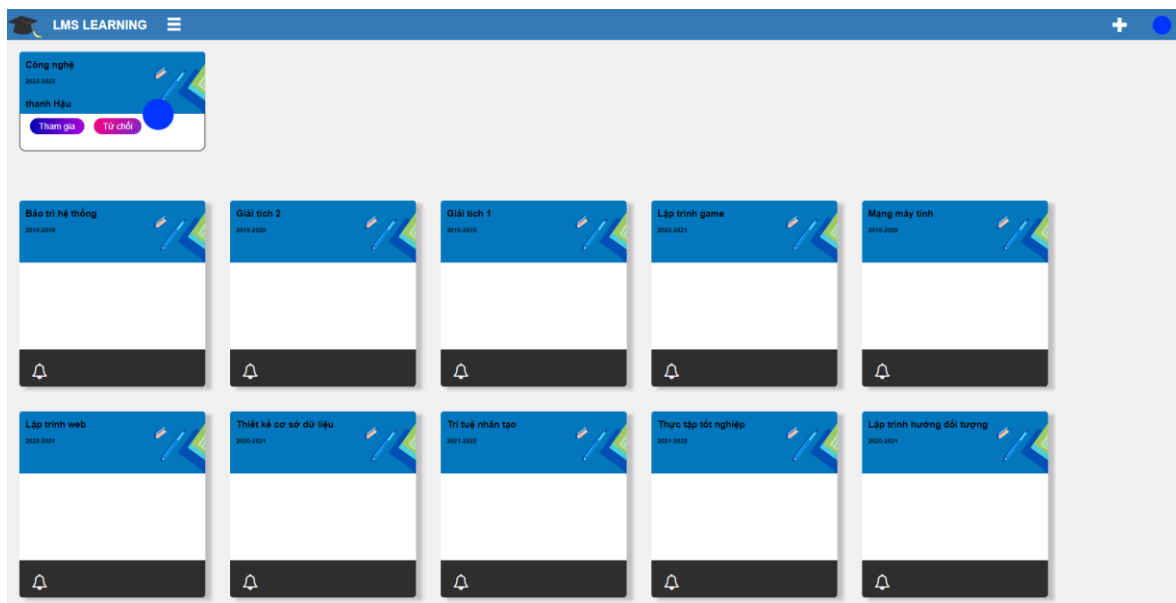


Hình 3. 2 Giao diện đăng nhập

Khi người dùng quên mật khẩu đăng nhập, người dùng có thể sử dụng chức năng quên mật khẩu. Người dùng cần nhập email và tên đăng nhập của tài khoản và sẽ có 1 mã xác nhận được gửi tới email. Nhập mã xác thực và điền thông tin mật khẩu mới.

3.2.2 Giao diện trang chủ

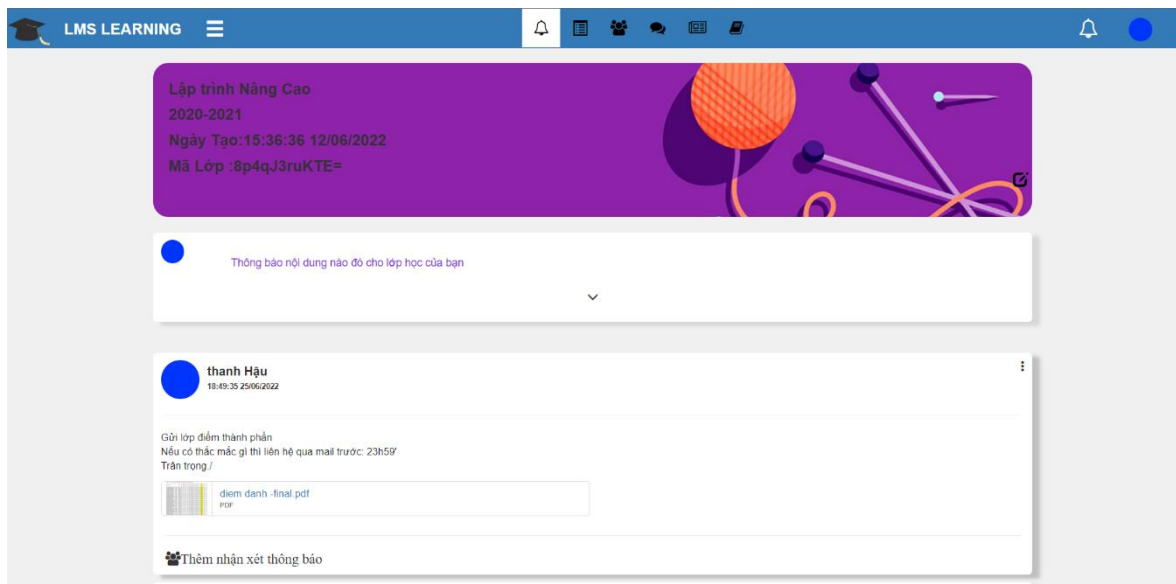
Giao diện trang chủ sẽ hiển thị lớp học người dùng đã tham gia và lời mời tham gia lớp học từ người khác. Giao diện sẽ hiển thị thông tin lớp từ ngày tạo mới nhất giúp người dùng theo dõi được các lớp mới nhất.



Hình 3. 3 Giao diện trang chủ

3.2.5 Giao diện thông báo

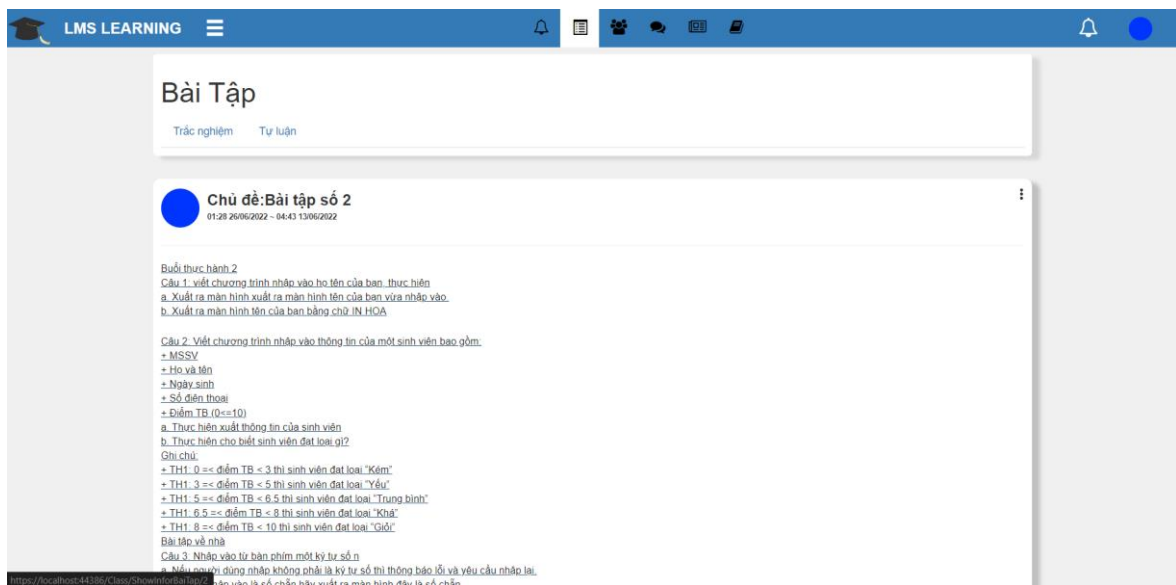
Giao diện hiển thị danh sách các thông báo của lớp học từ ngày tạo mới nhất. Giao diện sẽ hiển thị các thông tin của lớp học như tên lớp, chi tiết lớp, và nếu lớp học đó do người dùng tạo ra thì sẽ hiển thị thêm thông tin mã lớp để các thành viên khác có thể tham gia lớp học.



Hình 3. 4 Giao diện thông báo

3.2.7 Giao diện bài tập

Đối với giảng viên : Giao diện bài tập sẽ hiển thị danh sách bài tập giảng viên đã tạo.



Hình 3. 5 Giao diện bài tập cho giảng viên

Giảng viên chọn vào bài tập sẽ hiển thị thông tin điểm của từng sinh viên. Sẽ có hai giao diện khác nhau là giao diện điểm cho bài trắc nghiệm và giao diện điểm cho bài tự luận.

Giao diện điểm hiển thị xác thông tin của sinh viên như :

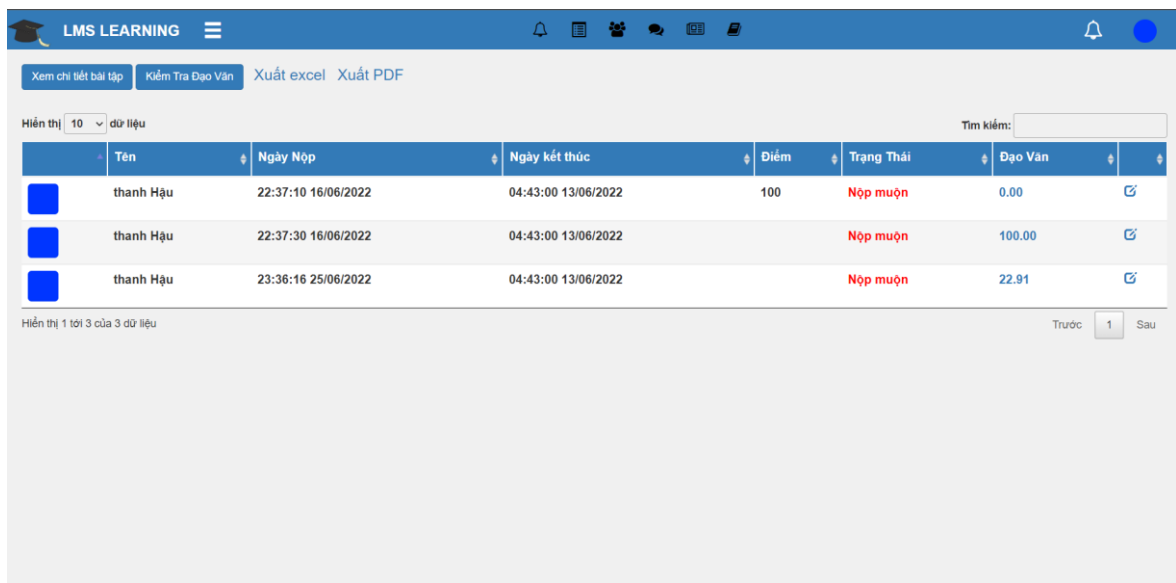
- Ảnh đại diện
- Họ tên của sinh viên
- Ngày nộp bài tập
- Ngày kết thúc
- Điểm bài tập
- Trạng thái : Hiển thị các trạng thái nộp bài như chưa nộp ,nộp muộn ,đã nộp
- Chi tiết bài tập sinh viên
- Phần trăm đạo văn

	Tên	Ngày Nộp	Ngày kết thúc	Điểm	Trạng Thái	
	thanh Hậu	18:03:13 25/06/2022	12:55:00 18/06/2022	56	Nộp muộn	
	thanh Hậu		12:55:00 18/06/2022		Chưa Nộp	
	thanh Hậu		12:55:00 18/06/2022		Chưa Nộp	

Hiện thị 1 tới 3 của 3 dữ liệu

Trước 1 Sau

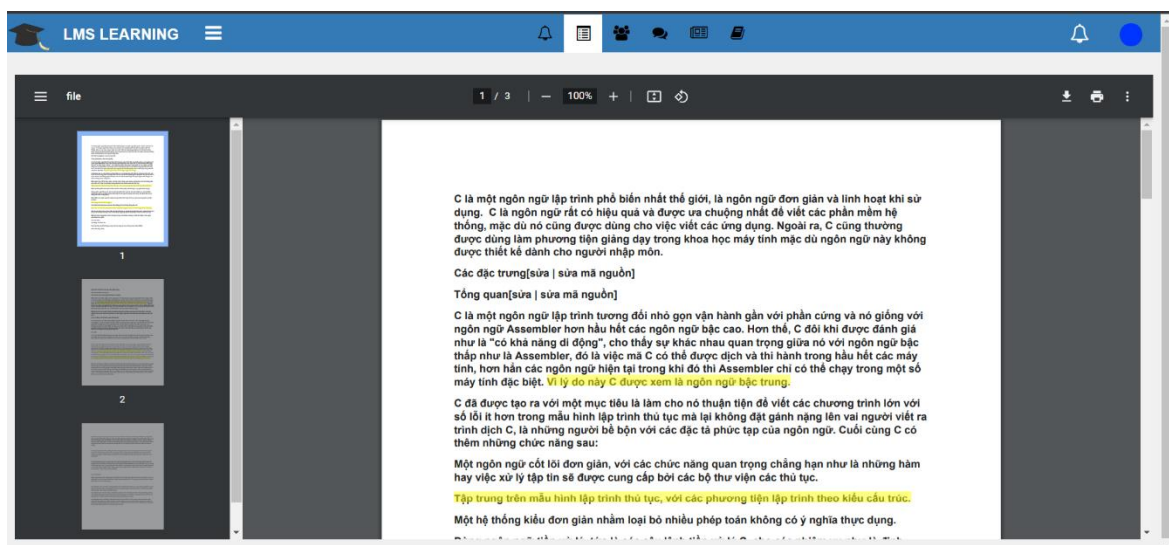
Hình 3. 6 Giao diện điểm bài tập trắc nghiệm



	Tên	Ngày Nộp	Ngày kết thúc	Điểm	Trạng Thái	Đạo Văn	
	thanh Hậu	22:37:10 16/06/2022	04:43:00 13/06/2022	100	Nộp muộn	0.00	
	thanh Hậu	22:37:30 16/06/2022	04:43:00 13/06/2022		Nộp muộn	100.00	
	thanh Hậu	23:36:16 25/06/2022	04:43:00 13/06/2022		Nộp muộn	22.91	

Hình 3. 7 Giao diện điểm bài tập tự luận

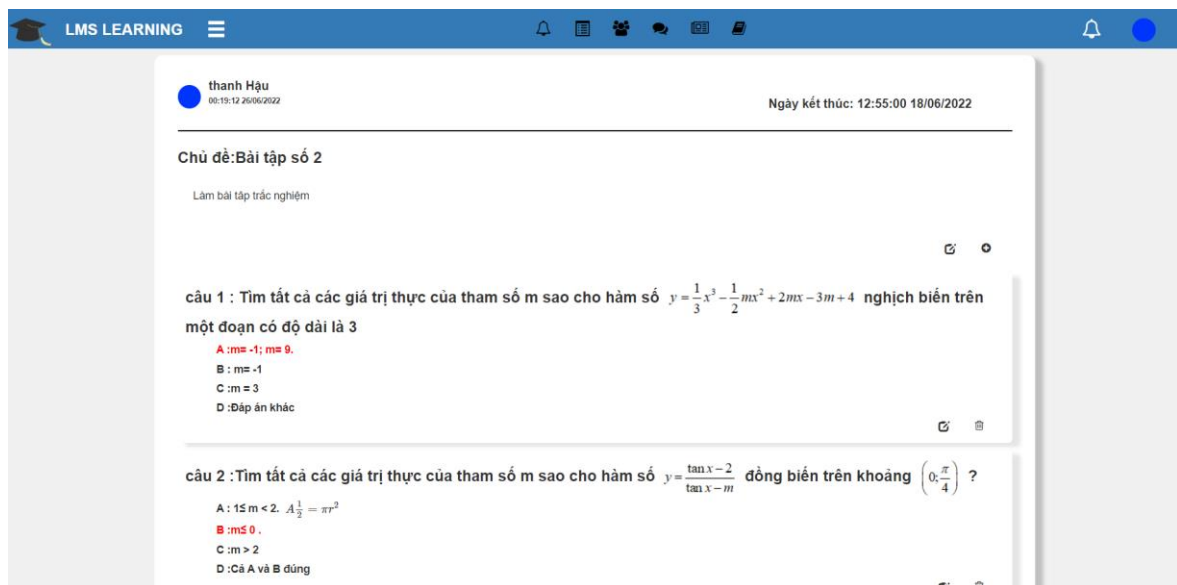
Đối với giao diện điểm bài tập tự luận giảng viên có thể kiểm tra đạo văn đối với bài tập của sinh viên. Giảng viên sẽ chọn từng sinh viên để xem so sánh giữa các bài làm.



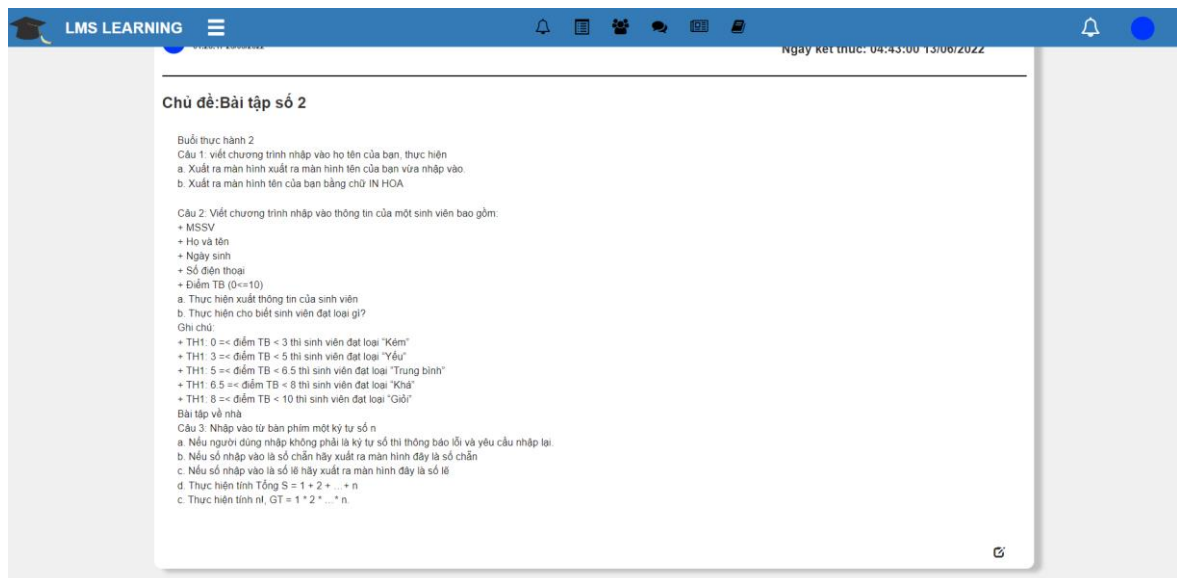
Hình 3. 8 Giao diện kiểm tra đạo văn

Trong giao diện điểm bài tập giảng viên có thể thay đổi thông tin của bài tập như :

- Thông tin bài tập như của đề bài tập ngày kết thúc của bài tập ...
- Thêm, sửa, xóa câu hỏi trắc nghiệm ...

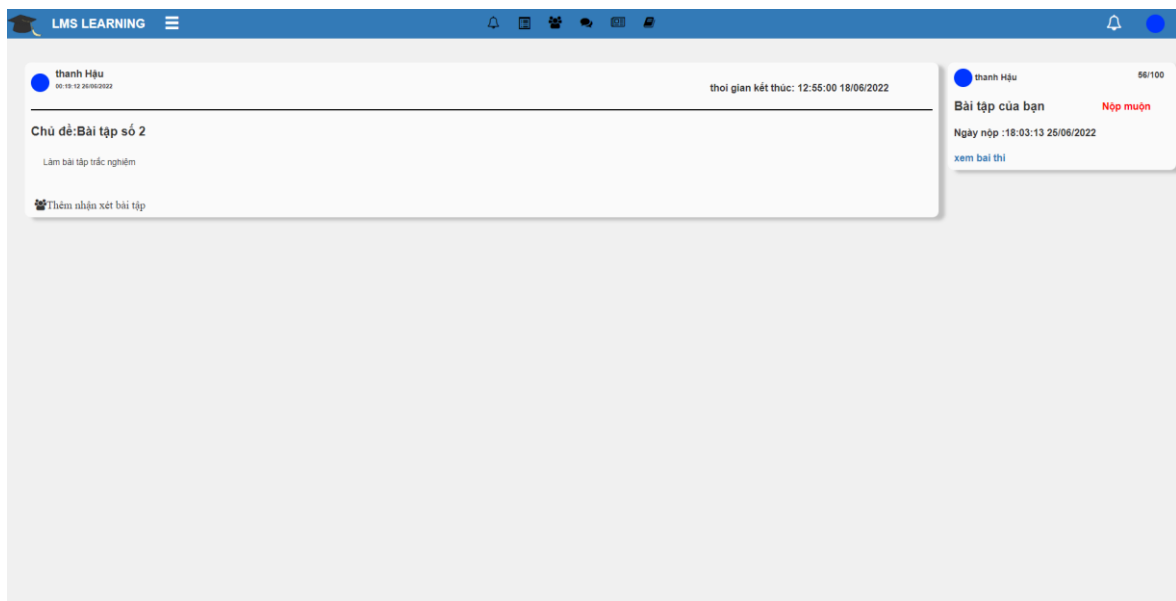


Hình 3. 9 Giao diện câu hỏi trắc nghiệm



Hình 3. 10 Giao diện câu hỏi tự luận

Ngoài các giao diện trên thì còn có giao diện thêm bài tập ,chấm điểm bài tập tự luận ,xem bài làm trắc nghiệm ...

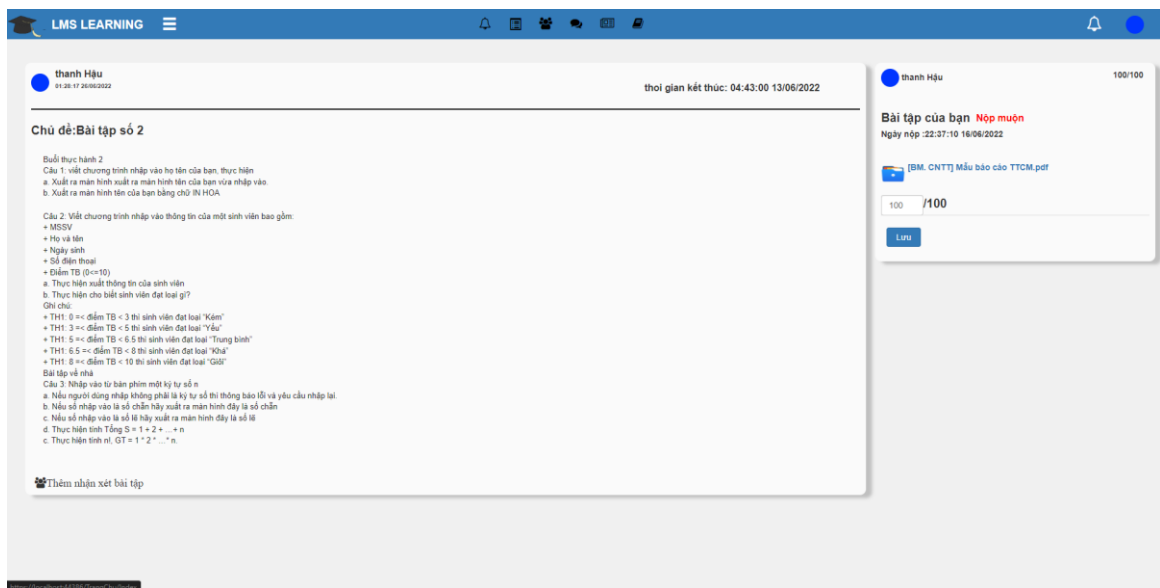


Hình 3. 11 Giao diện bài tập trắc nghiệm

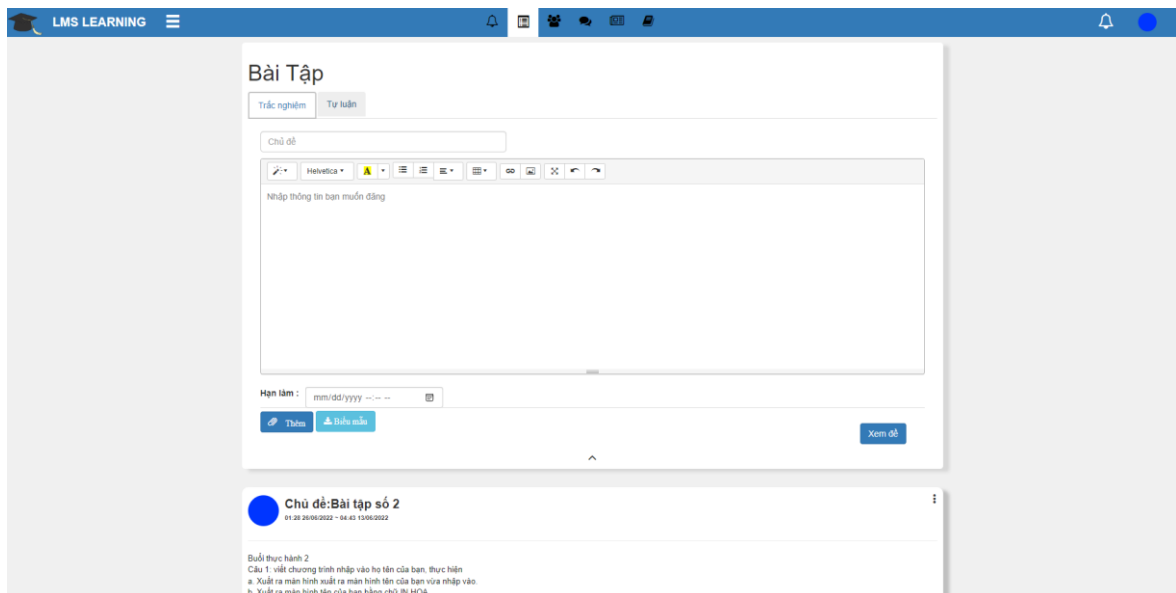


Hình 3. 12 Giao diện bài trắc nghiệm

Giao diện bài trắc nghiệm sẽ hiển thị đáp án đúng và đáp án sinh viên đã chọn

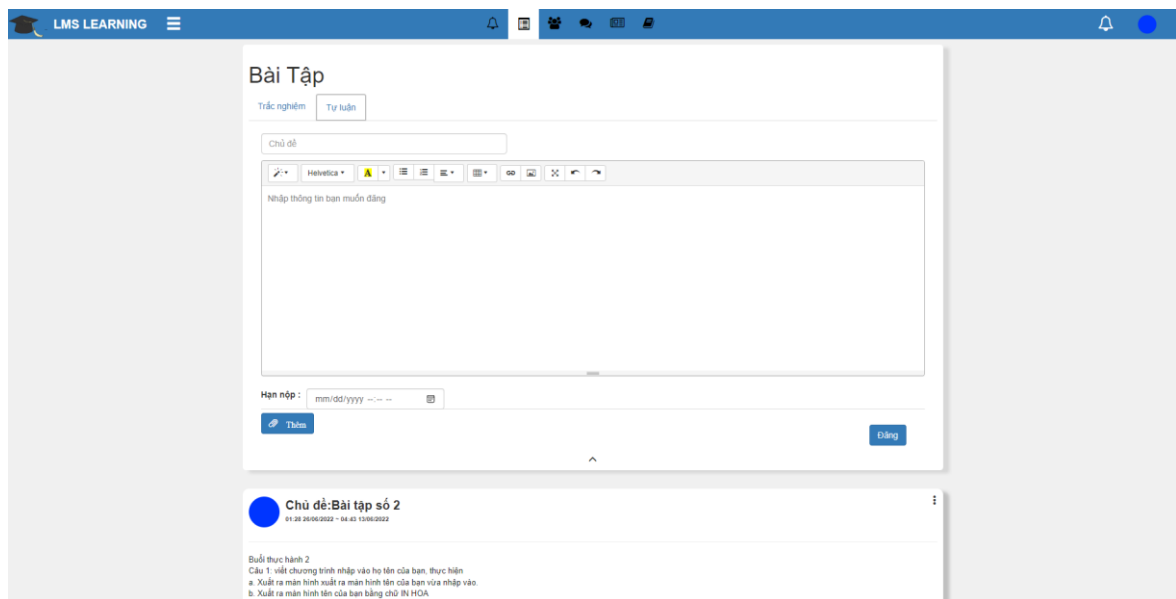


Hình 3. 13 Giao diện bài tập tự luận



Hình 3. 14 Giao diện tạo bài tập trắc nghiệm

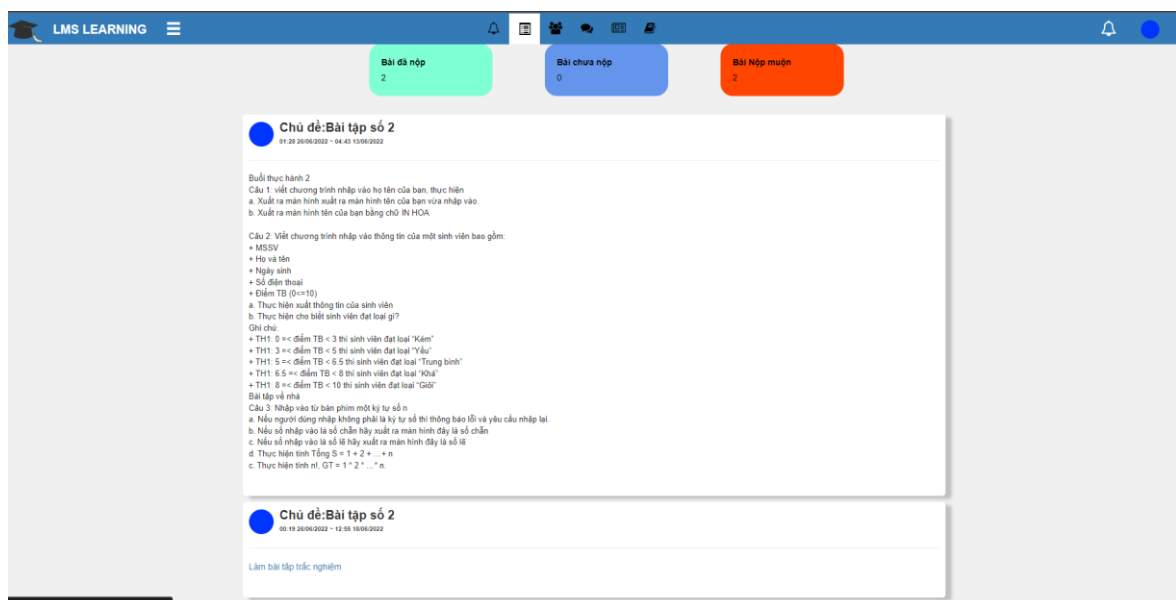
Giao diện bài tập trắc nghiệm yêu cầu giảng viên điền các thông tin như chủ đề bài tập, nội dung bài tập cùng với đó là file word câu hỏi trắc nghiệm theo format.



Hình 3. 15 Giao diện bài tập tự luận

Giao diện bài tập tự luận yêu cầu giảng viên điền các thông tin như chủ đề bài tập, nội dung bài tập cùng với đó là có thể có hoặc không có file liên quan tới bài tập tự luận.

Đối với sinh viên : Hiện thị các thông tin : Danh sách bài tập, số lượng bài tập nộp trễ ,chưa nộp, bài đã nộp

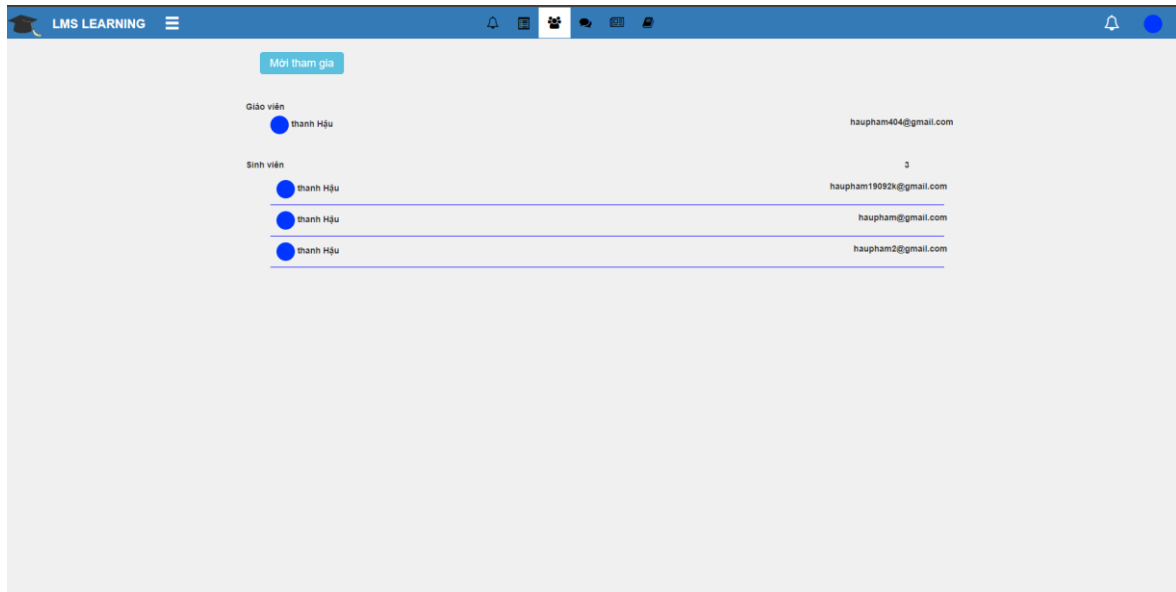


Hình 3. 16 Giao diện bài tập cho sinh viên

3.2.8 Giao diện thành viên lớp học

Giao diện hiển thị thông tin người tạo lớp học, số lượng sinh viên, thông tin mail ,họ tên của từng thành viên.

Với người dùng có vai trò là giảng viên : Giao diện sẽ hiển thị chức năng thêm thành viên



Hình 3. 17 Giao diện thành viên lớp học

3.2.9 Giao diện thống kê điểm

Xuất excel Xuất PDF

Hiện thị 10 dữ liệu

Tìm kiếm:

Hình đại diện	Họ tên	Email	Ngày Tham gia	Số lượng bài đã nộp	Số lượng bài trễ	Số lượng bài chưa nộp	Số lượng bài đã chấm	Điểm
	thanh Hậu	haupham19092k@gmail.com	15:38:28 12/06/2022	2	2	0	2	78
	thanh Hậu	haupham@gmail.com	01:00:16 14/06/2022	1	1	1	0	
	thanh Hậu	haupham2@gmail.com	23:35:53 25/06/2022	1	1	1	0	

Hiện thị 1 tới 3 của 3 dữ liệu

Trước 1 Sau

Hình 3. 18 Giao diện thống kê điểm

Giao diện thống kê điểm hiển thị các thông tin của sinh viên như :

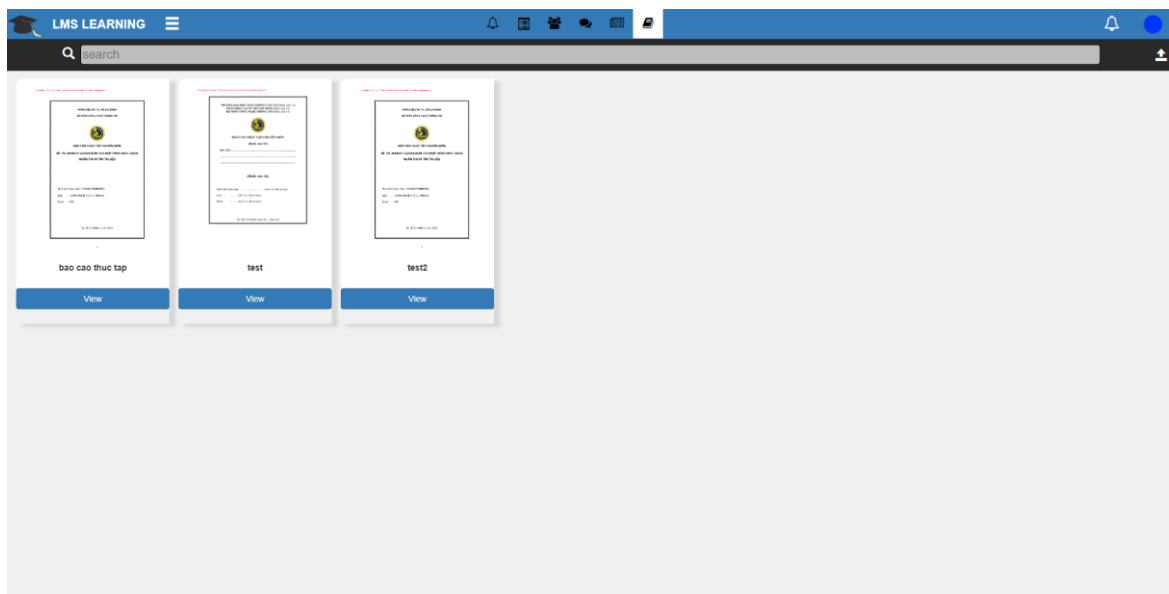
- Hình đại diện : Hiển thị hình ảnh sinh viên
- Họ tên
- Email
- Ngày tham gia
- Số lượng bài tập đã nộp

- Số lượng bài tập nộp trễ
- Số lượng bài tập chưa nộp
- Số lượng bài đã chấm điểm
- Điểm trung bình dựa vào điểm trung bình số lượng bài có điểm

3.2.10 Giao diện tài liệu

Hiện thị tất cả các tài liệu giảng viên lớp học đó đã chia sẻ cho sinh viên

Đối với giảng viên sẽ có thêm chức năng chia sẻ tài liệu



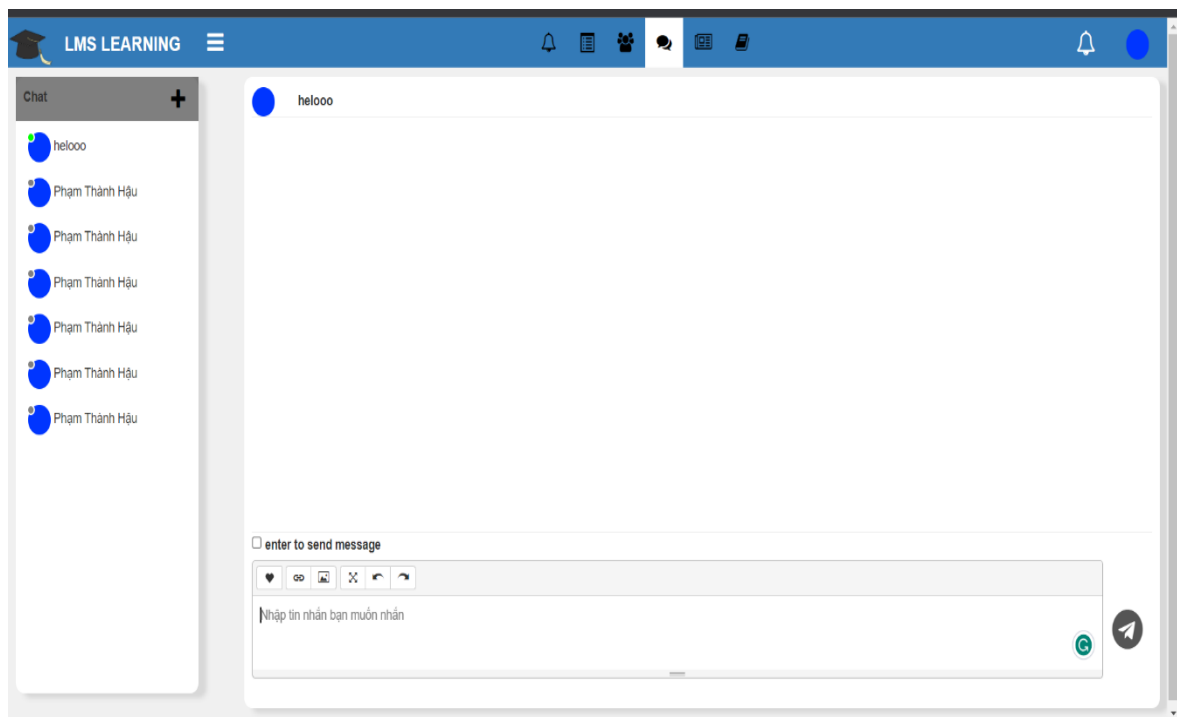
Hình 3. 19 Giao diện tài liệu



Hình 3. 20 Giao diện xem tài liệu

3.2.11 Giao diện nhắn tin

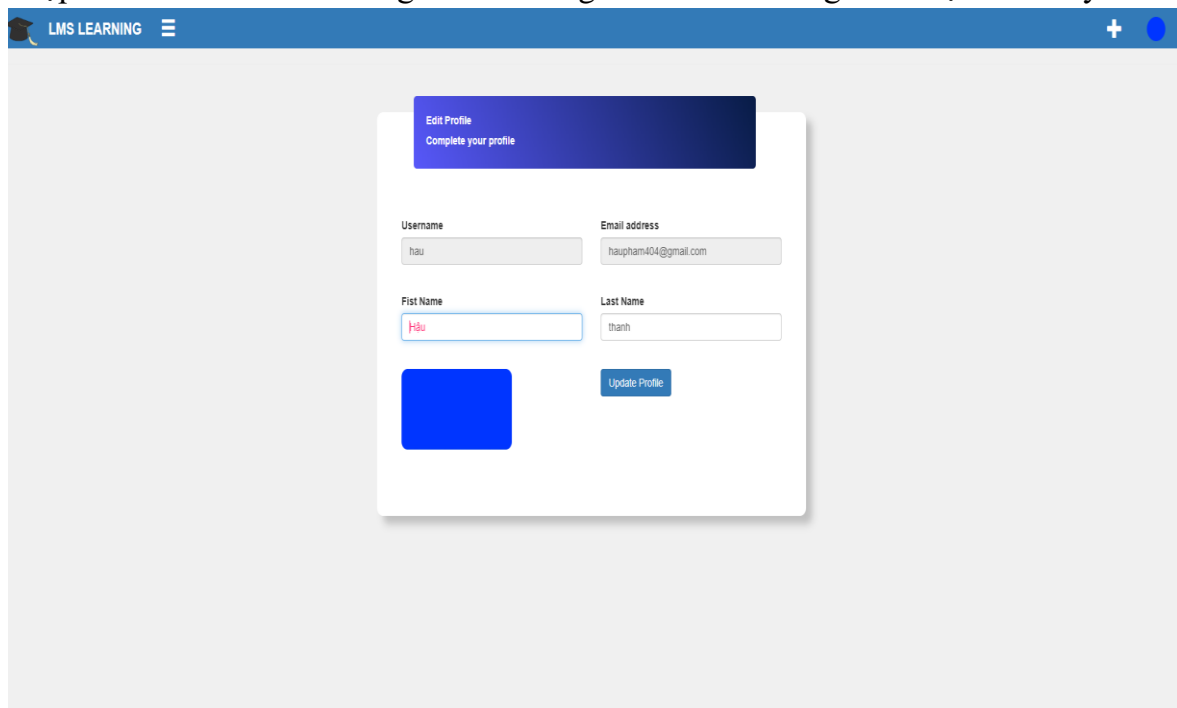
Giao diện hiển thị danh sách các thành viên trong lớp theo thứ tự tin nhắn mới nhất.



Hình 3. 21 Giao diện nhắn tin

3.2.12 Giao diện thông tin người dùng

Giao diện chỉ cho phép thay đổi họ, tên ,ảnh đại diện. Các thông tin như tên đăng nhập và email người dùng sẽ không được thay đổi.

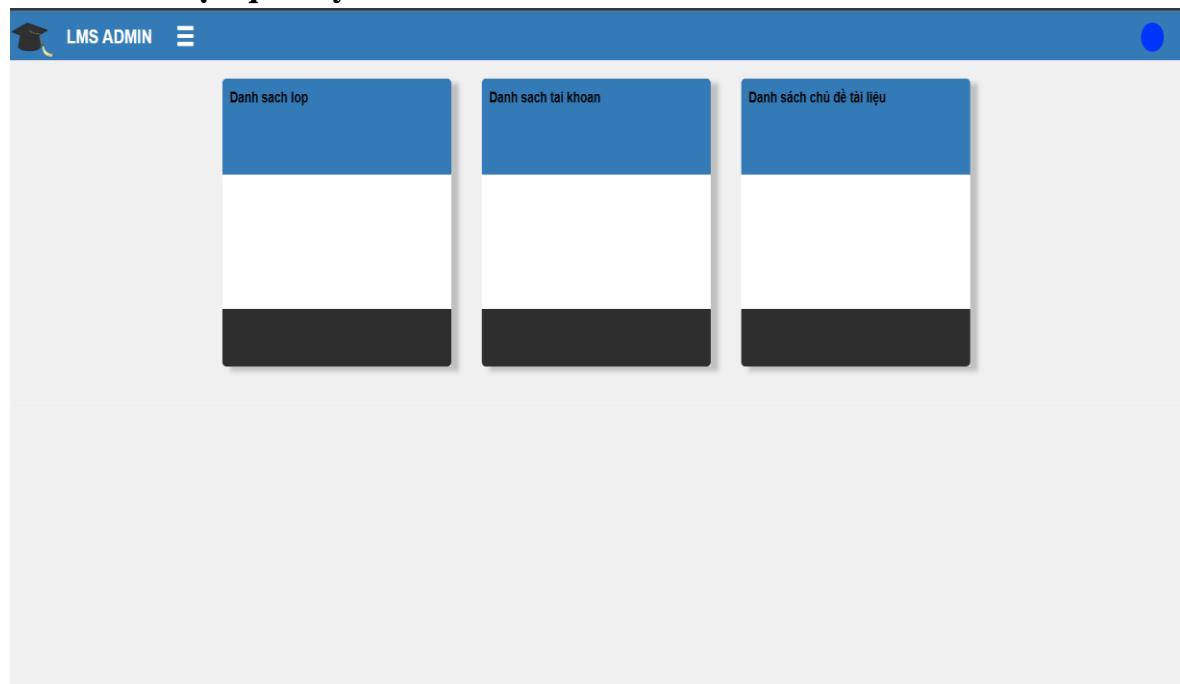


Hình 3. 22 Giao diện thông tin người dùng

Hình 3. 23 Giao diện thay đổi mật khẩu

Giao diện yêu cầu nhập mật khẩu cũ, mật khẩu mới và xác nhận mật khẩu mới

3.2.13 Giao diện quản lý của Admin



Hình 3. 24 Giao diện trang chủ Admin

Giao diện trang chủ dành cho admin gồm 3 chức năng :

- *Danh sách lớp học: Admin có thể xem thông tin lớp học và thành viên lớp học cũng như chuyển quyền của thành viên trong lớp học đó.*

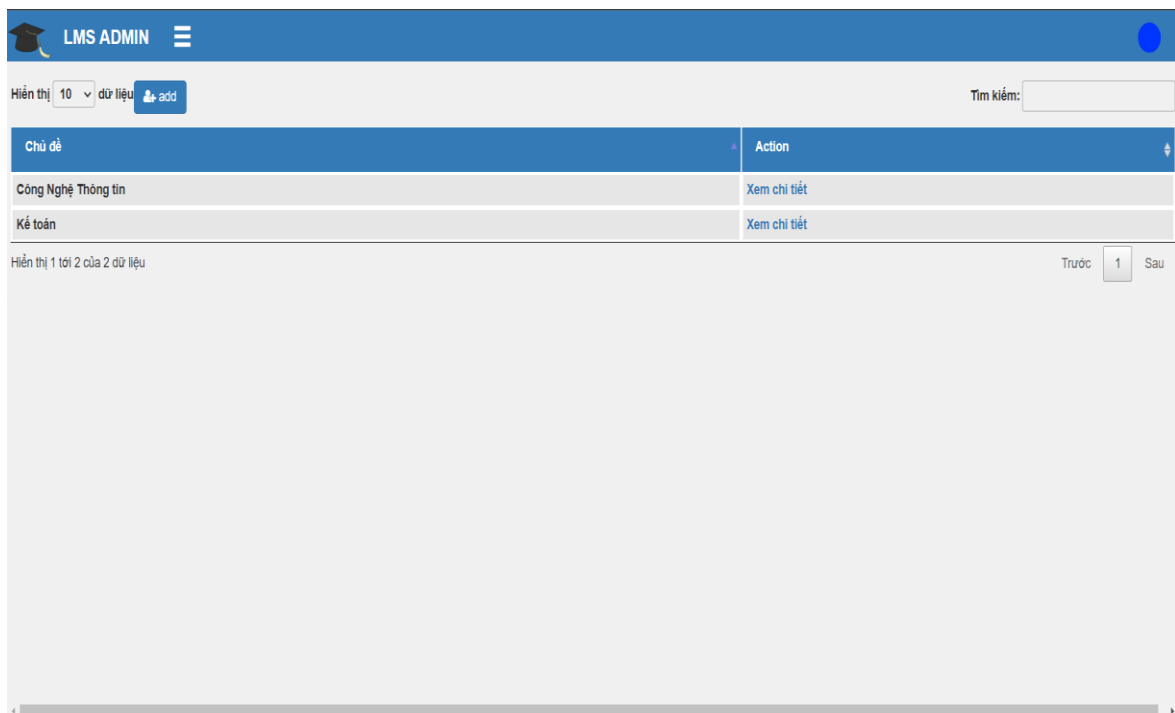
LMS ADMIN							
Hiện thị 10 dữ liệu					Tìm kiếm:		
Tên lớp học	Ngày tạo lớp học	Người tạo lớp học	Email người tạo lớp	Số lượng sinh viên	Số lượng giáo viên	Số lượng bài tập	Action
Test1	3/16/2023 4:47:41 PM	Phạm Thành Hậu	haupham404@gmail.com	1	6	2	Xem chi tiết
Hiện thị 1 tới 1 của 1 dữ liệu							
							Trước 1 Sau

Hình 3. 25 Giao diện quản lý lớp học dành cho Admin

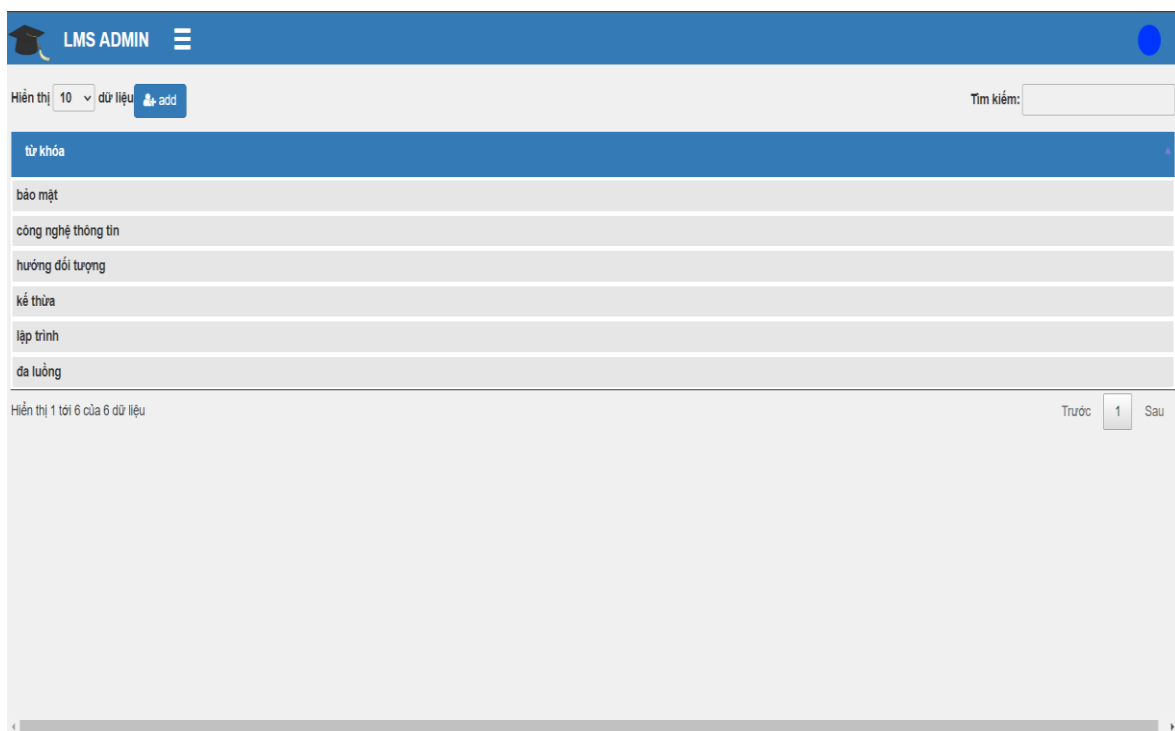
LMS ADMIN			
Hiện thị 10 dữ liệu		Tìm kiếm:	
Họ Tên	Chức vụ	Email	Hình đại diện
Phạm Thành Hậu	Sinh Viên	1@gmail.com	
Phạm Thành Hậu	Sinh Viên	2@gmail.com	
Phạm Thành Hậu	Sinh Viên	3@gmail.com	
Phạm Thành Hậu	Sinh Viên	4@gmail.com	
Phạm Thành Hậu	Sinh Viên	5@gmail.com	
Phạm Thành Hậu	Sinh Viên	6@gmail.com	
Phạm Thành Hậu	Giáo Viên	haupham404@gmail.com	
Hiện thị 1 tới 7 của 7 dữ liệu			
			Trước 1 Sau

Hình 3. 26 Giao diện quản lý thành viên lớp học dành cho Admin

- Danh sách chủ đề :dựa vào các tài liệu của người dung đã đưa lên hệ thống admin sẽ thêm các chủ đề tương ứng. Hệ thống sẽ dựa vào các chủ đề đó để gom cụm thành số cụm dữ liệu mong muốn













Hình 3. 27 Giao diện quản lý cụm dữ liệu



Hình 3. 28 Giao diện quản lý chi tiết cụm dữ liệu

- Giao diện quản lý tài khoản : Bao gồm tất cả tài khoản trong hệ thống. Ở chức năng này cho phép admin cấp quyền tạo lớp cho tài khoản.

LMS ADMIN			
Hiện thị 10 dữ liệu <button>add</button>		Tìm kiếm: <input type="text"/>	
Họ Tên	Chức vụ	Email	Hình đại diện
phạm thanh hau	Sinh Viên	10@gmail.com	
phạm thanh hau	Sinh Viên	11@gmail.com	
phạm thanh hau	Sinh Viên	12@gmail.com	
phạm thanh hau	Sinh Viên	13@gmail.com	
phạm thanh hau	Sinh Viên	14@gmail.com	
phạm thanh hau	Sinh Viên	15@gmail.com	
phạm thanh hau	Sinh Viên	16@gmail.com	
phạm thanh hau	Sinh Viên	17@gmail.com	
phạm thanh hau	Sinh Viên	18@gmail.com	
phạm thanh hau	Sinh Viên	19@gmail.com	
Hiện thị 1 tới 10 của 24 dữ liệu		Trước 1 2 3 Sau	

Hình 3. 29 Giao diện quản lý tài khoản

KẾT LUẬN VÀ KIẾN NGHỊ

1. Kết quả đạt được

Trong suốt khoảng thời gian nghiên cứu và thực hiện đề tài, em đã cố gắng làm hết khả năng của mình và vì thời gian làm đồ án có hạn nên kết quả đạt được vẫn còn nhiều hạn chế nhưng em cũng học hỏi được khá nhiều kiến thức mới mẻ cũng như chuyên sâu về thuật toán, lập trình... và những việc em đã đạt được như sau: Trình bày các kiến thức cơ bản về CSDL, phân tích thiết kế hệ thống, ngôn ngữ lập trình C#, golang. Tổng quan một số mô hình điển hình: mô hình thực thể liên kết, mô hình quan hệ, mô hình BFD, DFD các mức ngữ cảnh, đỉnh, dưới đỉnh. Tổng quan một số sơ đồ điển hình :sơ đồ tuần tự, sơ đồ hoạt động ,... Hiểu được quy trình phân tích và thiết kế một chương trình quản lý cơ bản. Biết được cách thức hoạt động của thuật toán kmeans trong thực tế cũng như cách sử lý ngôn ngữ.

2. Tồn tại

Trong quá trình làm đồ án, hiển nhiên sẽ có những lỗi chưa khắc phục hoàn toàn được nên sẽ có những chức năng chưa hoàn thiện. Số lượng chức năng chỉ ở mức tương đối như nhắn tin, thông báo ... Vì lý do thư viện cho office hiện chỉ hỗ trợ tối đa số trang nhỏ hơn 3 trang nên chức kiểm tra đạo văn hiện tại chỉ hỗ trợ cho tài liệu nhỏ hơn 3 trang cùng với đó đạo văn hiện chưa hỗ trợ kiểm tra về mặt ngữ nghĩa. Tập dữ liệu chạy thực nghiệm còn hạn chế về số lượng.

3. Hướng phát triển

Tiếp tục hoàn thiện và xây dựng thêm nhiều chức năng hỗ trợ trang web như: kiểm tra đạo văn về mặt ngữ nghĩa, phát triển thư viện cho office, sử dụng các kỹ thuật AI cho việc sử lý dữ liệu.

TÀI LIỆU THAM KHẢO

- [1] Professional ASP.NET MVC 5, J. Galloway.
- [2] Pro ASP.NET MVC 5 Platform, A. Freeman.
- [3] Programming Microsoft ASP.NET MVC, D. Esposito.
- [4] ASP.NET MVC-Building a Website with Visual Studio 2015 and Sharp:The Tactical Guidebook, J. Fagerberg, 2020.
- [5] Kỹ thuật lập trình Web với Ajax, P.T.Kiên, NXB Hồng Đức.
- [6] Lập trình C# từ cơ bản đến nâng cao, P.C.Ngô, 2007.
- [7] "Javascript", The JavaScript Workshop, J.L.E. a.J. Labrecque, 2019.
- [8] "JavaScript," in *The JavaScript Workshop*, J. L. E. a. Joseph Labrecque, 2020.
- [9] Learning Web Design: A Beginner's Guide to HTML, CSS, JavaScript, and Web Graphics, J. Robbins, 2020.
- [10] A new search algorithm for documents using blocks and words prefixes, Khalid Thabit and Sumaia M. AL-Ghuribi, 2013.
- [11] Plagiarism detection using document similarity based on distributed representation, Kensuke Baba, 2017.
- [12] Learning SignalR, Kenneth Yamikani, 2021.
- [13] Spire.Office for .NET Documentation, E-iceblue, 2021.
- [14] The Go Programming Language, Alan A. A. Donovan, Brian Kernighan, 2015.
- [15] <https://www.techtarget.com/searchcio/definition/learning-management-system>, truy cập 11/06/2023, Kate Brush, 2019.
- [16] Echo Web Framework Documentation, Echo ,2021
- [17] <https://www.freecodecamp.org/news/cron-jobs-in-linux>, truy cập 11/06/2023, Zaira Hira, 2021
- [18] Database System Concepts, Abraham Silberschatz, Henry F. Korth, S. Sudarshan, 2010

- [19] <https://getbootstrap.com/docs/4.1/getting-started/introduction>, truy cập 11/06/2023, Bootstrap, 2020
- [20] Introduction to Data Mining, Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar, 2019.