

**ĐẠI HỌC QUỐC GIA HÀ NỘI  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**



**BÁO CÁO BÀI TẬP LỚN CUỐI KỲ BỘ MÔN TƯỞNG TÁC  
NGƯỜI MÁY  
NĂM HỌC 2024 – 2025**

**NHÓM 1**

Giảng viên hướng dẫn

Môn học

Lớp học phần

Thành viên

TS. Ngô Thị Duyên

Tương tác người máy

2425I INT2041 57

Lê Thế Hiển - 22028101

Lê Văn Đức - 22028041

Phạm Quý Sơn - 22028296

Nguyễn Quang Khải - 22028299

**HÀ NỘI - 2024**

# Mục lục

<b>1</b>	<b>Giới thiệu bài toán</b>	<b>4</b>
1.1	Bối cảnh và động lực . . . . .	4
1.2	Mục tiêu của dự án . . . . .	4
1.3	Phạm vi và giới hạn của dự án . . . . .	4
<b>2</b>	<b>Tổng quan và Cơ sở lý thuyết</b>	<b>5</b>
2.1	Tổng quan về ngôn ngữ ký hiệu . . . . .	5
2.2	Các phương pháp học ngôn ngữ ký hiệu hiện nay . . . . .	5
2.3	Nền tảng về Tương tác Người-Máy (HMI) . . . . .	6
<b>3</b>	<b>Phân tích và Thiết kế hệ thống</b>	<b>6</b>
3.0.1	Nhóm đối tượng người dùng . . . . .	6
3.1	Mô tả chức năng của ứng dụng . . . . .	6
3.2	Thiết kế giao diện và trải nghiệm người dùng (Phi chức năng) . . . . .	7
<b>4</b>	<b>Thiết kế hệ thống</b>	<b>7</b>
4.1	Kiến trúc và Cấu trúc tổng quan hệ thống . . . . .	7
4.1.1	Frontend (Giao diện người dùng): . . . . .	7
4.1.2	Backend (Xử lý logic): . . . . .	7
4.1.3	Database . . . . .	7
4.1.4	Các services khác . . . . .	8
4.2	Đặc tả hệ thống và Usecase . . . . .	8
4.2.1	Use Case 1: Đăng ký . . . . .	10
4.2.2	Use Case 2: Đăng nhập . . . . .	12
4.2.3	Use Case 3: Học tập ký hiệu thủ ngữ . . . . .	13
4.2.4	Use Case 4:Thực hành và kiểm tra . . . . .	14
4.2.5	Use Case 5: Phiên dịch . . . . .	16
4.2.6	Use Case 6: Từ Điển . . . . .	17
4.2.7	Use Case 7: Kiểm tra ký hiệu . . . . .	18
4.2.8	Use Case 8: Theo dõi bảng xếp hạng . . . . .	20
4.2.9	Use Case 9: Cộng đồng . . . . .	21
4.3	Góc nhìn các lớp tham gia . . . . .	22
4.3.1	Chức năng 1: Kiểm tra ký hiệu . . . . .	22
4.3.2	Chức năng 2: Khoá học . . . . .	23
4.3.3	Chức năng 3: Ôn tập . . . . .	23
4.3.4	Chức năng 4: Phiên dịch . . . . .	23
4.3.5	Chức năng 5: Từ điển . . . . .	24
4.3.6	Chức năng 6: Thống kê và xếp hạng . . . . .	24

4.4	Thiết kế cơ sở dữ liệu . . . . .	26
4.4.1	Users (Người dùng): . . . . .	26
4.4.2	learningdata (Bài học): . . . . .	26
4.4.3	revisedata (Kết quả thực hành): . . . . .	26
4.4.4	Libbase_data (Cộng đồng): . . . . .	27
4.4.5	Community_data (Cộng đồng): . . . . .	27
4.4.6	document_data (Cộng đồng): . . . . .	27
4.4.7	Social_data (Cộng đồng): . . . . .	27
4.4.8	Statistic_data (Thống kê): . . . . .	28
4.4.9	Dictionary_data (Từ điển): . . . . .	28
4.5	Thiết kế hệ thống thông minh . . . . .	28
4.5.1	Hệ thống nhận diện ngôn ngữ kí hiệu . . . . .	28
4.5.2	Hệ thống dịch ngôn ngữ kí hiệu . . . . .	30
<b>5</b>	<b>Phát triển và Công nghệ sử dụng</b>	<b>31</b>
5.1	Công nghệ và công cụ phát triển . . . . .	31
5.2	Quy trình phát triển . . . . .	32
5.3	Các thách thức kỹ thuật . . . . .	32
5.4	Phương pháp kiểm thử ứng dụng . . . . .	32
<b>6</b>	<b>Kết quả và Đánh giá</b>	<b>33</b>
6.1	Đánh giá hiệu quả của ứng dụng . . . . .	33
6.2	Các điểm cải tiến và tiềm năng mở rộng . . . . .	33
<b>7</b>	<b>Thành Tựu</b>	<b>33</b>
7.1	Tóm tắt thành tựu đạt được . . . . .	33
7.2	Hướng phát triển trong tương lai . . . . .	33
<b>8</b>	<b>Bảng phân chia công việc</b>	<b>34</b>
<b>9</b>	<b>Phụ lục và Tài liệu tham khảo</b>	<b>34</b>
9.1	Phụ lục . . . . .	34
9.2	Tài liệu tham khảo . . . . .	35

# 1 Giới thiệu bài toán

## 1.1 Bối cảnh và động lực

Ngôn ngữ ký hiệu hay còn gọi là thủ ngữ là phương tiện giao tiếp vô cùng quan trọng đối với cộng đồng người điếc và khiếm thính. Không những thế, đối với những người đối diện, người nghe, những người gặp những vấn đề về nói, cũng như những người muốn cảm thông, quan tâm và giao tiếp với họ. Thủ ngữ cũng là một con đường có thể giúp đỡ và gắn kết mọi người với nhau hơn.

Tuy nhiên, việc học tập ngôn ngữ ký hiệu vẫn gặp nhiều trở ngại do thiếu sách giáo khoa chất lượng, phương pháp dạy hạn chế và sự hỗ trợ hạn chế từ công nghệ thông tin. Với sự phát triển không ngừng của công nghệ di động và trí tuệ nhân tạo, việc áp dụng các công nghệ này vào việc học ngôn ngữ ký hiệu mang lại nhiều cơ hội để cải thiện trải nghiệm học tập cho người điếc.

Trên cơ sở đó, SignLearn ra đời với mục đích hỗ trợ người dùng, giúp đỡ phần nào trong việc khó khăn về tiếp cận nội dung học tập, tài liệu liên quan, cộng đồng – xã hội cùng chung hoàn cảnh và điều kiện.

## 1.2 Mục tiêu của dự án

Dự án hướng tới xây dựng một trang web hỗ trợ và giúp đỡ mọi người trong việc dạy học, và học tập để có thể giao tiếp dễ dàng hơn với người khiếm thính thông qua các khía cạnh:

- **Phát triển website** hỗ trợ học ngôn ngữ ký hiệu cho người câm điếc, người có nhu cầu học ký hiệu thủ ngữ cho mục đích giao tiếp.
- **Cung cấp các tài liệu học tập**, nguồn tài liệu phong phú bao gồm video, hình ảnh và các bài báo, trang web cộng đồng hữu ích
- **Theo dõi tiến độ**, giúp những người liên quan theo dõi tiến độ người dùng và đưa ra phản hồi cá nhân hóa phù hợp.
- **Tạo ra một cộng đồng hỗ trợ** nơi người dùng có thể nhận được sự giúp đỡ từ các chuyên gia và người dùng khác.

## 1.3 Phạm vi và giới hạn của dự án

**Phạm vi:** Ứng dụng website tập trung vào việc học ngôn ngữ ký hiệu Việt Nam.

**Giới hạn:** Dự án chưa bao gồm việc phát triển ứng dụng cho phiên bản cho máy tính bảng hoặc thiết bị đeo thông minh. Chỉ hỗ trợ cho việc sử dụng trang web ở các layout cho thiết bị di động và máy tính bảng. Chưa có đột phá về mặt trí tuệ nhân tạo, chỉ nằm ở mức phân tích cử chỉ, ký hiệu và dự đoán kết quả dựa trên dữ liệu đã training.

## 2 Tổng quan và Cơ sở lý thuyết

### 2.1 Tổng quan về ngôn ngữ ký hiệu

Ngôn ngữ ký hiệu, hay còn gọi là thủ ngữ, là một hệ thống giao tiếp phức tạp sử dụng các cử chỉ tay, biểu cảm khuôn mặt, tư thế cơ thể và các yếu tố thị giác khác để truyền đạt thông tin và ý nghĩa. Đây không chỉ đơn thuần là những cử động tay mà còn bao gồm các quy tắc ngữ pháp, cú pháp và từ vựng riêng biệt, cho phép người sử dụng diễn đạt các ý tưởng phức tạp tương đương với ngôn ngữ nói thông thường.

Mỗi quốc gia hay vùng lãnh thổ đều phát triển hệ ngôn ngữ ký hiệu riêng, phản ánh nét văn hóa, truyền thống và đặc điểm ngôn ngữ bản địa của địa phương đó. Ví dụ, Ngôn ngữ Ký hiệu Việt Nam (VSVN) có những đặc trưng riêng biệt, khác với các ngôn ngữ ký hiệu khác như American Sign Language (ASL) tại Mỹ hay British Sign Language (BSL) tại Anh. VSVN không chỉ đơn thuần là sự chuyển dịch từ tiếng nói sang cử chỉ mà còn mang trong mình những giá trị văn hóa và xã hội đặc thù của người Việt.

VSVN được sử dụng rộng rãi trong cộng đồng người điếc tại Việt Nam, đóng vai trò quan trọng trong việc giao tiếp hàng ngày, giáo dục, và hòa nhập xã hội cho người khiếm thính. Việc sử dụng ngôn ngữ ký hiệu không chỉ giúp nâng cao khả năng giao tiếp mà còn góp phần vào việc bảo tồn và phát triển bản sắc văn hóa của cộng đồng người điếc. Ngoài ra, ngôn ngữ ký hiệu còn hỗ trợ trong các lĩnh vực như giáo dục đặc biệt, y tế và các hoạt động xã hội, tạo điều kiện thuận lợi cho người điếc tiếp cận thông tin và tham gia vào các hoạt động cộng đồng một cách đầy đủ và hiệu quả. Các phần mềm kể trên đều hỗ trợ người khiếm thị ở nhiều khía cạnh khác nhau của cuộc sống bằng cách cung cấp các nhóm chức năng khác nhau như: tìm đồ vật, đọc văn bản, v.v. Tuy nhiên, nhóm nhận định rằng các sản phẩm tuy ứng dụng rất thiết thực và có ích cho người khiếm thị, tuy nhiên chưa thực sự có giao diện thân thiện - dễ sử dụng với đối tượng mà họ hướng tới.

### 2.2 Các phương pháp học ngôn ngữ ký hiệu hiện nay

Hiện nay, có nhiều phương pháp học ngôn ngữ ký hiệu như:

- **Học truyền thống:** Tham gia các lớp học trực tiếp với giáo viên.
- **Tài liệu học tập:** Sử dụng sách, video và ứng dụng di động.
- **Luyện tập giao tiếp:** Tham gia các nhóm giao tiếp hoặc cộng đồng người điếc.
- **Sử dụng công nghệ:** Các ứng dụng, phần mềm hỗ trợ học ký hiệu thông qua tương tác và phản hồi tức thì.

## 2.3 Nền tảng về Tương tác Người-Máy (HMI)

Tương tác Người-Máy (Human-Machine Interaction - HMI) là lĩnh vực nghiên cứu đa ngành tập trung vào việc thiết kế, đánh giá và triển khai các giao diện người dùng hiệu quả, dễ sử dụng và thân thiện. HCI không chỉ nghiên cứu về cách con người tương tác với máy tính mà còn xem xét cách các hệ thống máy tính có thể hỗ trợ và nâng cao trải nghiệm người dùng.

Trong bối cảnh phát triển ứng dụng "Sign Learning", HMI đóng vai trò quan trọng trong việc đảm bảo rằng ứng dụng không chỉ đáp ứng được các yêu cầu chức năng mà còn mang lại trải nghiệm người dùng tối ưu cho người điếc và người khiếm thính. Điều này bao gồm việc thiết kế giao diện người dùng (UI) trực quan, sử dụng các biểu tượng và màu sắc dễ nhận biết, và cung cấp các phản hồi trực quan thay vì âm thanh.

## 3 Phân tích và Thiết kế hệ thống

### 3.0.1 Nhóm đối tượng người dùng

- **Người dùng chính:** Người điếc và người khiếm thính, người mất khả năng nói hoặc giao tiếp, người có nhu cầu trong việc học tập ký hiệu thủ ngữ.
- **Người dùng liên quan:**
  - **Giáo viên và chuyên gia:** Giáo viên, người hướng dẫn, người mất khả năng nói hoặc giao tiếp, người có nhu cầu trong việc giảng dạy, học tập ký hiệu thủ ngữ.
  - **Người thân và người hỗ trợ:** Có thể sử dụng ứng dụng để hỗ trợ việc học của người điếc.

### 3.1 Mô tả chức năng của ứng dụng

- **Học tập ký hiệu qua video và hình ảnh:** Ứng dụng cung cấp kho nội dung đa dạng bao gồm video hướng dẫn và hình ảnh minh họa các cử chỉ ký hiệu, giúp người học dễ dàng nắm bắt và ghi nhớ.
- **Thực hành ký hiệu và hoạt động tương tác:** Các bài tập thực hành cho phép người dùng ghi lại cử chỉ của mình và so sánh với mẫu chuẩn, đồng thời tham gia hoạt động ôn tập nhằm ghi nhớ kiến thức.
- **Theo dõi tiến bộ:** Ứng dụng cung cấp bảng điều khiển cá nhân cho phép người dùng theo dõi tiến độ học tập, điểm số và các mục tiêu đã đạt được.

- **Hỗ trợ từ chuyên gia:** Tính năng này cho phép người dùng đặt câu hỏi và nhận hỗ trợ từ các chuyên gia ngôn ngữ ký hiệu thông qua hệ thống chat hoặc cuộc gọi video.

### 3.2 Thiết kế giao diện và trải nghiệm người dùng (Phi chức năng)

- **Mẫu giao diện (wireframes):** Giao diện của ứng dụng được thiết kế đơn giản, trực quan với các biểu tượng rõ ràng và dễ hiểu. Các màn hình chính bao gồm trang chủ, kho tài liệu học, bài tập thực hành và trang cá nhân. . .
- **Thiết kế khả năng tiếp cận:** Ứng dụng tuân thủ các nguyên tắc thiết kế khả năng tiếp cận, đảm bảo rằng tất cả các tính năng đều có thể sử dụng được bởi người dùng có nhu cầu đặc biệt, như hỗ trợ phóng to màn hình thu nhỏ màn hình, thích hợp với nhiều thiết bị và giao diện thân thiện với người điếc.

## 4 Thiết kế hệ thống

### 4.1 Kiến trúc và Cấu trúc tổng quan hệ thống

Hệ thống được chia làm ba lớp chính

#### 4.1.1 Frontend (Giao diện người dùng):

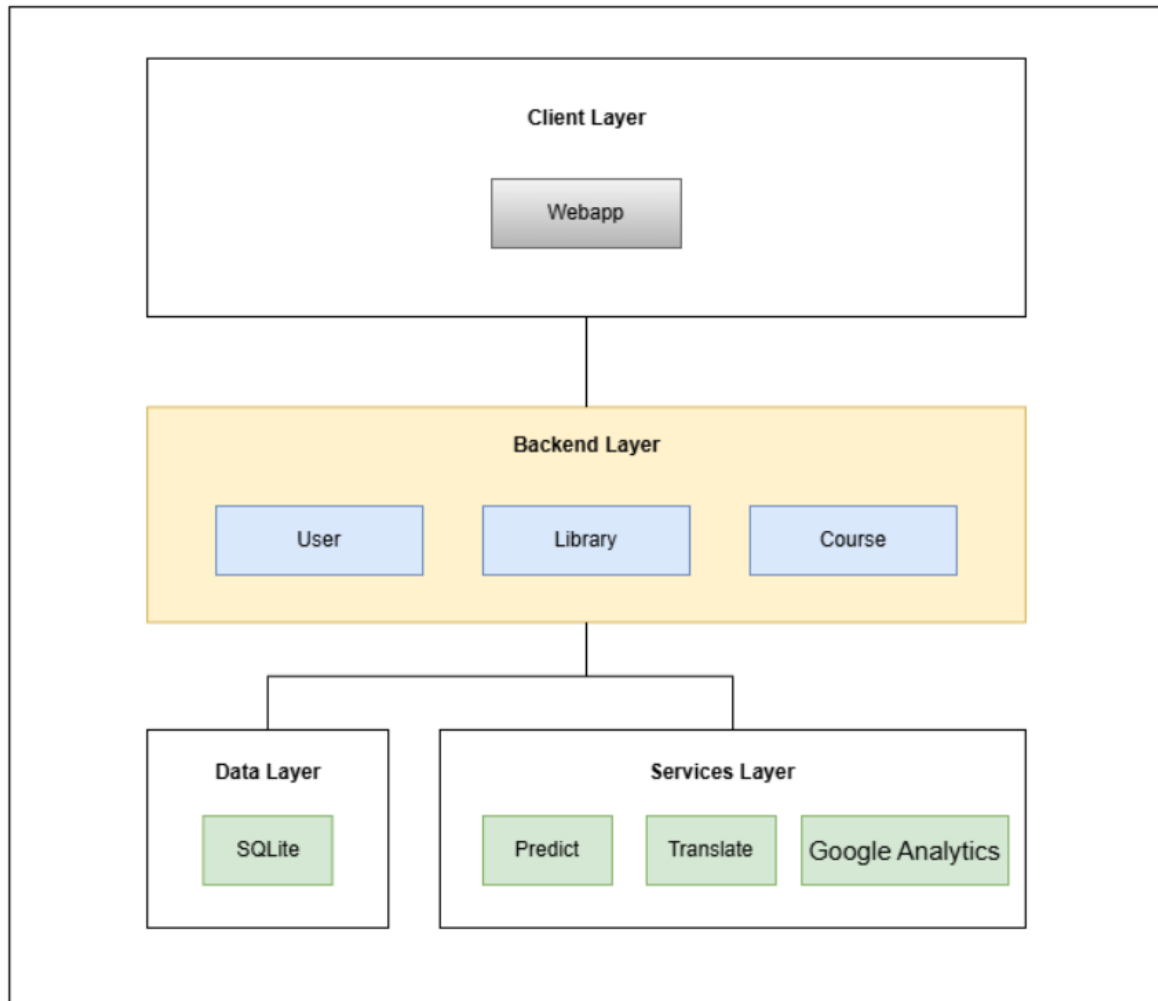
- Sử dụng framework Next.js kết hợp thư viện UI Mantine để thiết kế giao diện thân thiện, dễ sử dụng.
- Giao diện được tối ưu hóa để hỗ trợ cả trên thiết bị di động và máy tính.
- Cung cấp các tính năng chính như học ngôn ngữ ký hiệu qua video, thực hành, và theo dõi tiến bộ.

#### 4.1.2 Backend (Xử lý logic):

- Dựa trên framework Flask để quản lý yêu cầu và xử lý dữ liệu từ frontend.
- Các API được tối ưu để hỗ trợ xử lý nhanh chóng các yêu cầu như lưu trữ kết quả học tập, phân tích dữ liệu, và nhận diện cử chỉ.

#### 4.1.3 Database

- Sử dụng SQLite để demo hệ thống
- Nâng cấp PostgreSQL quan hệ để xử lý các liên kết của dữ liệu



#### 4.1.4 Các services khác

- Sử dụng Google Analytics để phân tích hành vi người dùng.
- Các mô hình học sâu (CNN kết hợp MobileNetV2 và GRU) xử lý nhận diện và phân loại cử chỉ từ video.

## 4.2 Đặc tả hệ thống và Usecase

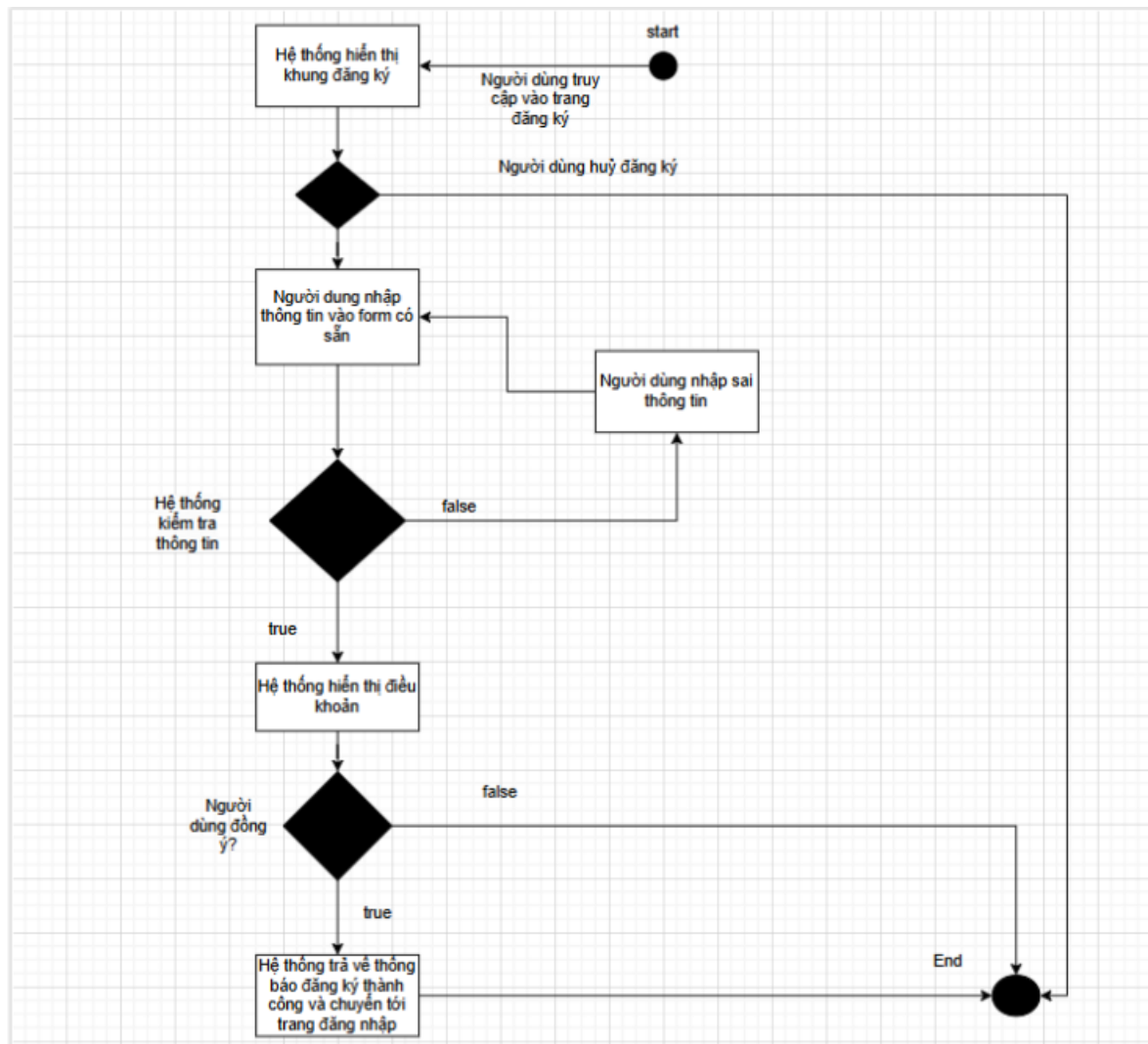
Hệ thống tập trung vào các usecase sau:

- **Đăng ký:** Người dùng tạo mới tài khoản để sử dụng với hệ thống
- **Đăng nhập:** Người dùng đăng nhập để sử dụng hệ thống
- **Học tập ký hiệu thủ ngữ:** Người dùng vào trang khoá học, chọn video, thực hiện theo hướng dẫn thủ ngữ trên video



- **Thực hành và kiểm tra:**Người dùng vào trang ôn tập, chọn video, thực hiện theo hướng dẫn thủ ngữ trên video, đoán ra từ ngữ chính xác với nội dung video đã mô tả
- **Phiên dịch:** Người dùng đến trang khoá học, chọn mục phiên dịch, chọn video tiếng việt mà mình có, hoặc nhập từ tiếng việt, gửi tới hệ thống, hệ thống phân tích và trả về video tiếng anh tương ứng
- **Từ điển:** Người dùng đến trang khoá học, chọn mục từ điển, tìm kiếm từ vựng mà mình muốn học
- **Kiểm tra ký hiệu:** Người dùng đến trang khoá học, vào mục kiểm tra ký hiệu, thực hiện quay video và nhận kết quả từ model
- **Theo dõi bảng xếp hạng:**Hiển thị danh sách bảng xếp hạng người dùng
- **Theo dõi tiến trình học tập:**Người dùng đến trang khoá học, chọn mục thống kê, ở đây thống kê lại số từ, số chữ và số giờ đã học trong thời gian theo tuần
- **Cộng đồng:**Người dùng đến trang tài liệu, lựa chọn thông tin mình cần tìm hiểu, cộng đồng để trao đổi

#### 4.2.1 Use Case 1: Đăng ký



**Mô tả:** Use case cho phép đăng ký tài khoản mới để sử dụng ứng dụng.

**Các tác nhân**

- Người dùng.
- Hệ thống.

Điều kiện kích hoạt ca sử dụng:

- Người dùng muốn tạo tài khoản mới để kết nối với hệ thống

Tiền điều kiện:

- Người dùng chưa đăng ký tài khoản trước đó
- Người dùng có đủ thông tin cần thiết.
- Thiết bị của người dùng được kết nối Internet.

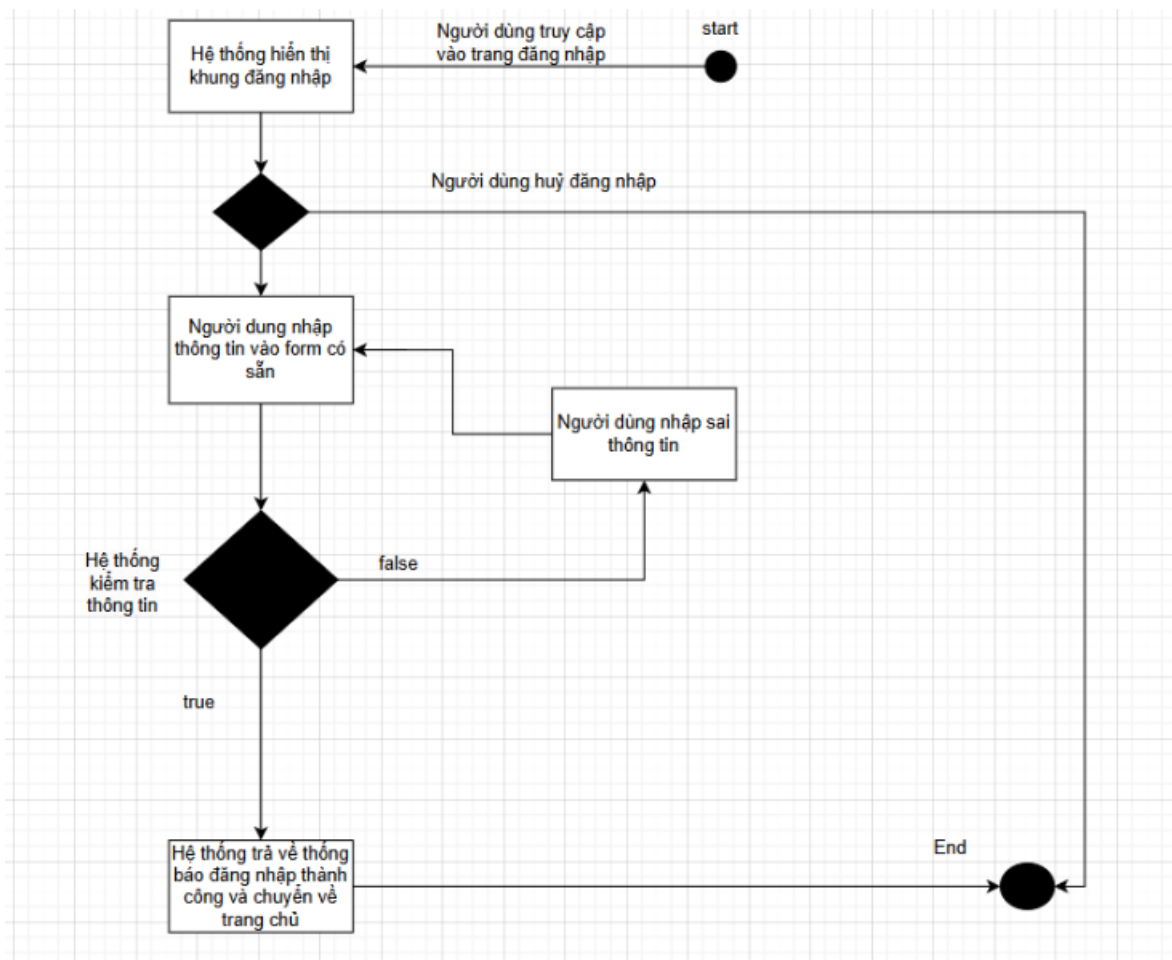
**Hậu điều kiện:**

- Người dùng đăng ký thành công.
- Thông tin người dùng được lưu vào cơ sở dữ liệu người dùng.

**Các luồng sự kiện:**

- Luồng cơ bản:
  - Người dùng truy cập vào trang web
  - Người dùng chọn đăng ký tài khoản
  - Người dùng nhập thông tin cá nhân
  - Hệ thống nhận thông tin, kiểm tra và trả về hướng dẫn
  - Hệ thống thông báo đăng ký thành công và trả về trang đăng nhập
- Luồng ngoại lệ:
  - Hệ thống không thể hoàn tất do lỗi kết nối Internet.
  - Người dùng kiểm tra lại kết nối và thử tải lại.
  - Người dùng nhập sai thông tin
  - Người dùng kiểm tra và nhập lại thông tin

#### 4.2.2 Use Case 2: Đăng nhập



**Mô tả:** Use case cho phép người dùng đăng nhập để sử dụng hệ thống  
**Các tác nhân**

- Người dùng.
- Hệ thống.

**Điều kiện kích hoạt ca sử dụng:**

- Người dùng muốn đăng nhập để sử dụng hệ thống

**Tiền điều kiện:**

- Người dùng đã có tài khoản trước đó
- Thiết bị của người dùng được kết nối Internet.

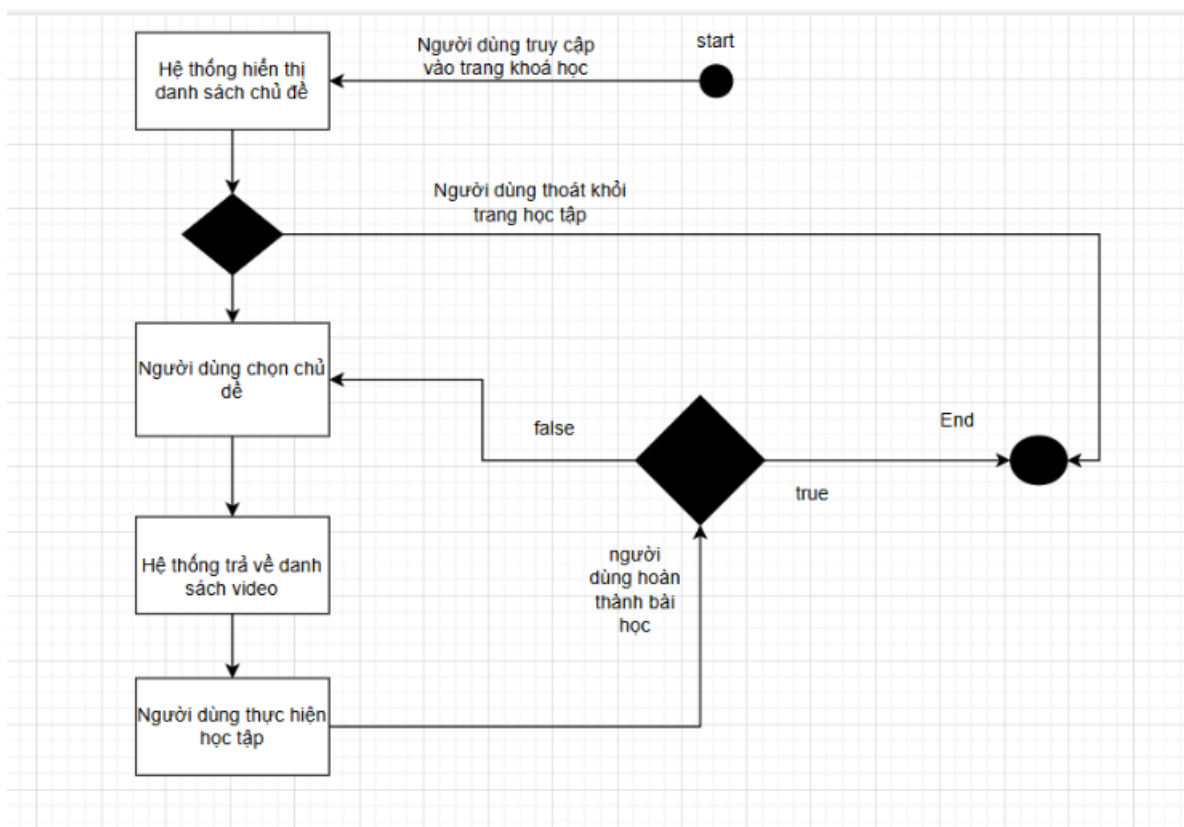
**Hậu điều kiện:**

- Người dùng đăng nhập thành công, hệ thống trả về trang chủ

**Các luồng sự kiện:**

- Luồng cơ bản:
  - Người dùng truy cập vào trang web
  - Người dùng chọn đăng nhập tài khoản
  - Người dùng nhập thông tin cá nhân
  - Hệ thống nhận thông tin, kiểm tra và trả về trang chủ
- Luồng ngoại lệ:
  - Hệ thống không thể hoàn tất do lỗi kết nối Internet.
    - \* Người dùng kiểm tra lại kết nối và thử tải lại.
  - Người dùng nhập sai thông tin
    - \* Người dùng kiểm tra và nhập lại thông tin

#### 4.2.3 Use Case 3: Học tập ký hiệu thủ ngữ



#### Các tác nhân

- Người dùng.
- Hệ thống.

Điều kiện kích hoạt ca sử dụng:

- Người dùng muốn học tập ký hiệu thủ ngữ

#### Tiền điều kiện:

- Thiết bị của người dùng hỗ trợ Internet.

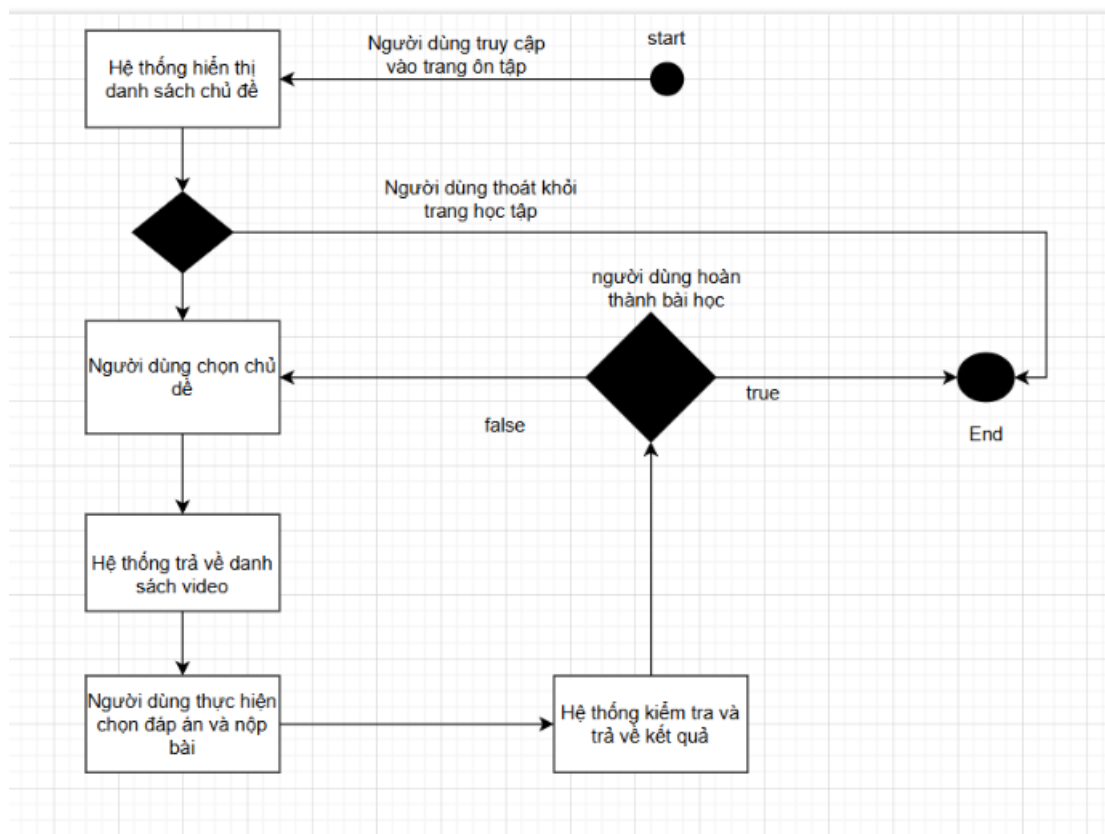
#### Hậu điều kiện:

- Tiến trình hoàn thành của phần học được cập nhật

#### Các luồng sự kiện:

- Luồng cơ bản:
  - Người dùng truy cập vào trang khoá học.
  - Hệ thống hiển thị danh sách chủ đề
  - Người dùng chọn chủ đề muốn học
  - Hệ thống trả về danh sách video phát video và yêu cầu người dùng thực hành theo.
- Luồng ngoại lệ:
  - Hệ thống không hoạt động do không có kết nối internet
  - Người dùng thực hiện lại bài tập.

#### 4.2.4 Use Case 4: Thực hành và kiểm tra



### **Các tác nhân**

- Người dùng.
- Hệ thống.

### **Điều kiện kích hoạt ca sử dụng:**

- Người dùng muốn kiểm tra khả năng nhận diện và thực hành thủ ngữ.

### **Tiền điều kiện:**

- Thiết bị của người dùng hỗ trợ Internet.

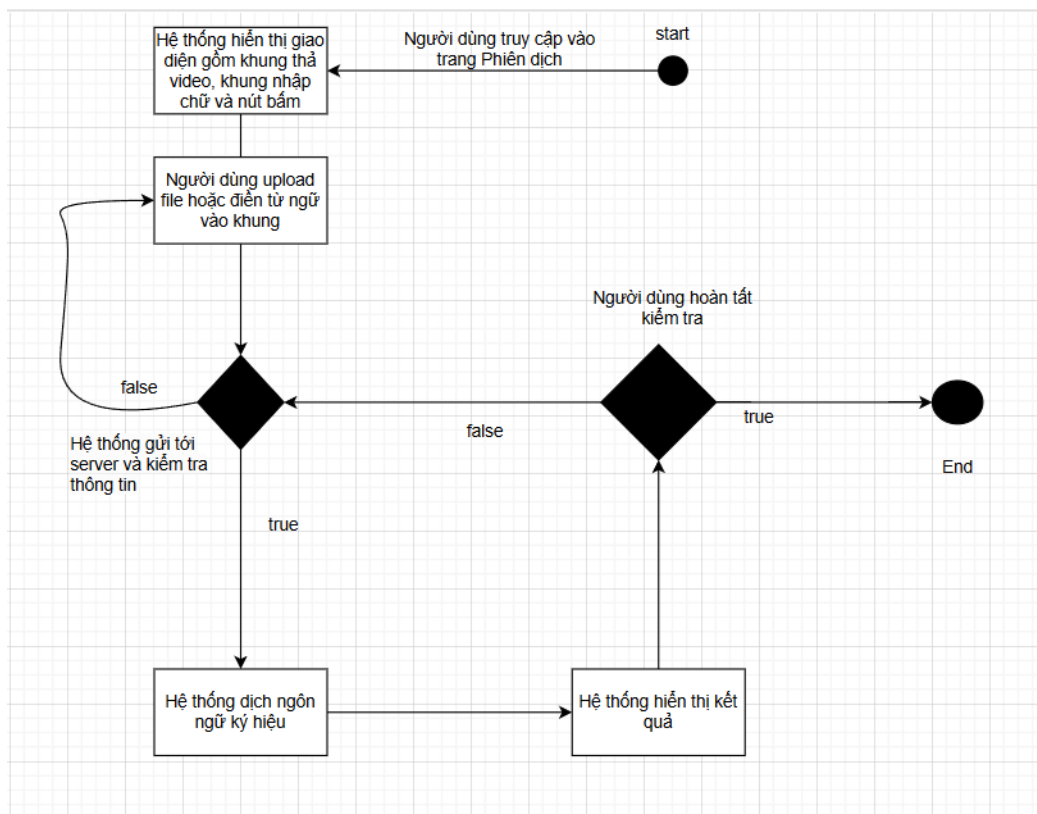
### **Hậu điều kiện:**

- Điểm số của người dùng được lưu vào cơ sở dữ liệu.
- Tiến trình thực hành được cập nhật.

### **Các luồng sự kiện:**

- Luồng cơ bản:
  - Người dùng truy cập vào trang Ôn tập.
  - Hệ thống hiển thị danh sách chủ đề ôn tập.
  - Người dùng chọn một chủ đề.
  - Hệ thống phát video và yêu cầu người chọn kết quả chính xác
  - Người dùng chọn đáp án
  - Hệ thống kiểm tra kết quả
  - Hệ thống hiển thị kết quả và gợi ý cải thiện.
- Luồng ngoại lệ:
  - Hệ thống không hoạt động do không có kết nối internet
  - Người dùng thực hiện lại bài tập.

#### 4.2.5 Use Case 5: Phiên dịch



**Mô tả:** Người dùng muốn thực hiện phiên dịch video từ tiếng việt sang tiếng anh.

**Các tác nhân**

- Người dùng.
- Hệ thống.

**Điều kiện kích hoạt ca sử dụng:**

- Người dùng muốn thực hành thủ ngữ.

**Tiền điều kiện:**

- Thiết bị của người dùng hỗ trợ ghi hình và kết nối Internet.

**Hậu điều kiện:**

- Kết quả được trả về

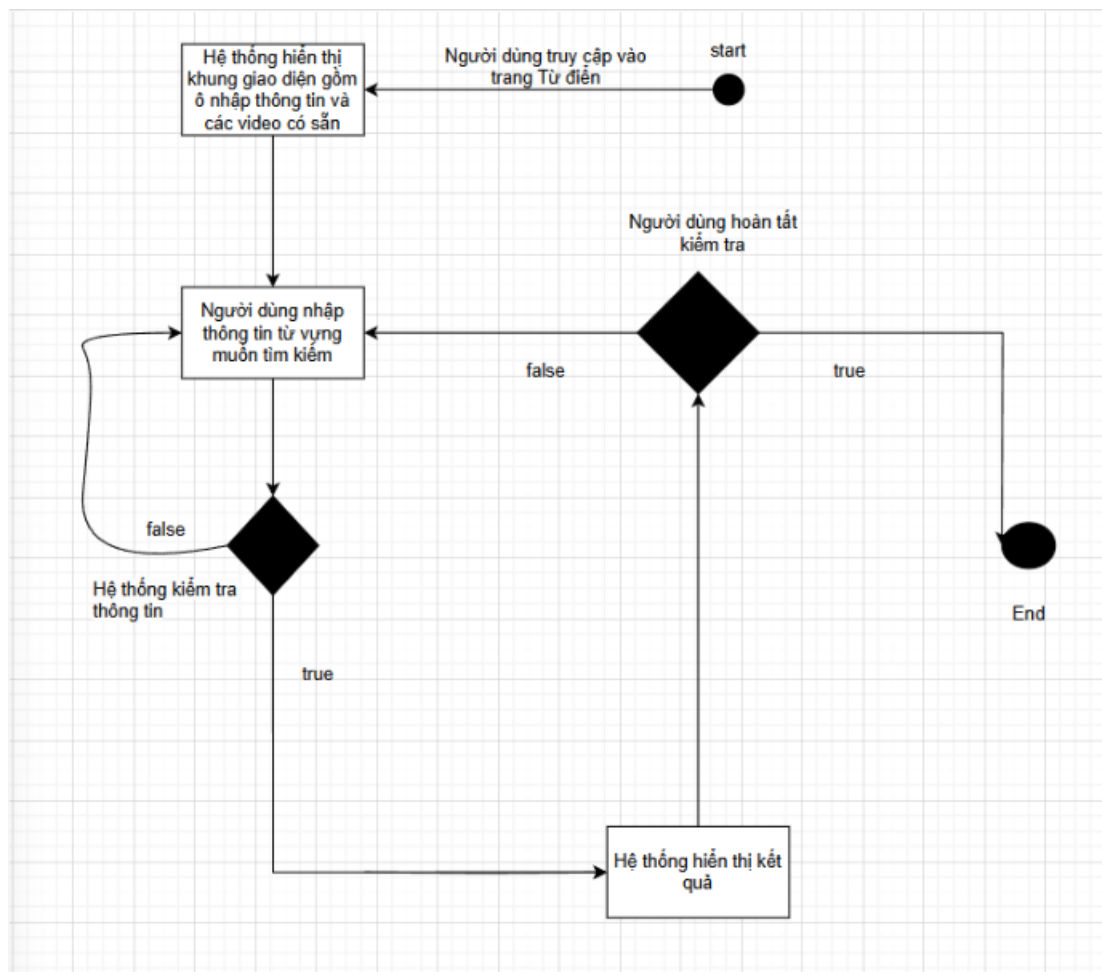
**Các luồng sự kiện:**

- Luồng cơ bản:
  - Người dùng truy cập vào trang khoá học, chọn mục phiên dịch.



- Người dùng thực hiện upload video tiếng việt hoặc nhập từ ngữ tiếng việt mà mình muốn phiên dịch thành video tiếng anh
  - Hệ thống kiểm tra dự đoán và trả về kết quả
  - Hệ thống hiển thị kết quả
- Luồng ngoại lệ:
    - Hệ thống không nhận diện được video hoặc từ ngữ
    - Người dùng thực hiện lại thao tác

#### 4.2.6 Use Case 6: Từ Điển



**Mô tả:** Người dùng muốn thực hiện tra cứu từ vựng.

**Các tác nhân**

- Người dùng.
- Hệ thống.

**Điều kiện kích hoạt ca sử dụng:**

- Người dùng muốn thực hiện tra cứu từ vựng.

#### Tiền điều kiện:

- Thiết bị của người dùng hỗ trợ kết nối Internet.

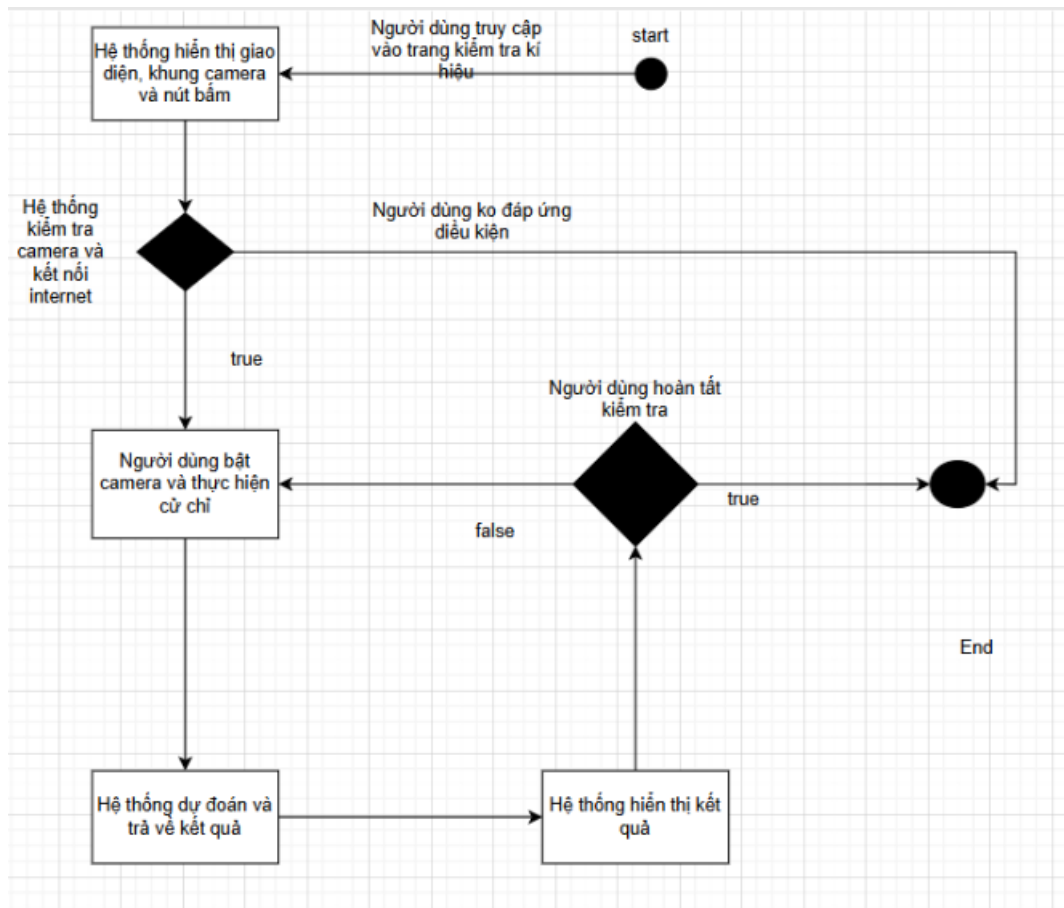
#### Hậu điều kiện:

- Kết quả được trả về

#### Các luồng sự kiện:

- Luồng cơ bản:
  - Người dùng truy cập vào trang khoá học, chọn mục Từ điển.
  - Người dùng nhập từ vựng vào ô tìm kiếm
  - Hệ thống kiểm tra và trả về kết quả
  - Hệ thống hiển thị kết quả
- Luồng ngoại lệ:
  - Hệ thống không tìm thấy từ vựng.
  - Người dùng thực hiện ghi lại

#### 4.2.7 Use Case 7: Kiểm tra ký hiệu



### **Các tác nhân**

- Người dùng.
- Hệ thống.

### **Điều kiện kích hoạt ca sử dụng:**

- Người dùng muốn thực hành thủ ngữ.

### **Tiền điều kiện:**

- Thiết bị của người dùng hỗ trợ ghi hình và kết nối Internet.

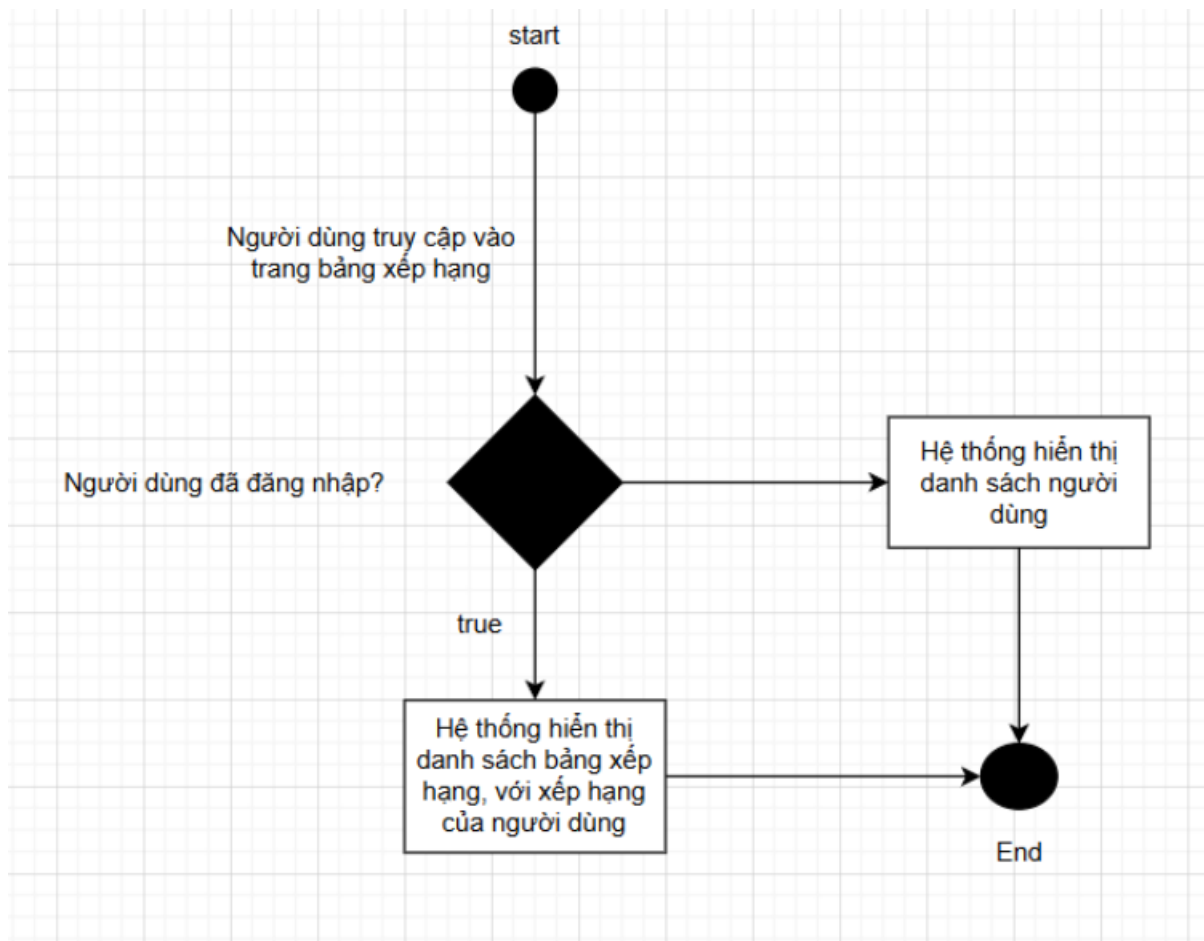
### **Hậu điều kiện:**

- Kết quả được trả về

### **Các luồng sự kiện:**

- Luồng cơ bản:
  - Người dùng truy cập vào trang khoá học, chọn mục kiểm tra ký hiệu.
  - Người dùng thực hiện các ký hiệu và hệ thống ghi lại cử chỉ.
  - Hệ thống kiểm tra dự đoán và trả về kết quả
  - Hệ thống hiển thị kết quả
- Luồng ngoại lệ:
  - Hệ thống không nhận diện được cử chỉ.
  - Người dùng thực hiện ghi lại

#### 4.2.8 Use Case 8: Theo dõi bảng xếp hạng



##### Các tác nhân

- Người dùng.
- Hệ thống.

##### Điều kiện kích hoạt ca sử dụng:

- Người dùng muốn biết vị trí của mình so với người khác.

##### Tiền điều kiện:

- Người dùng đã hoàn thành ít nhất một bài học hoặc bài thực hành.

##### Hậu điều kiện:

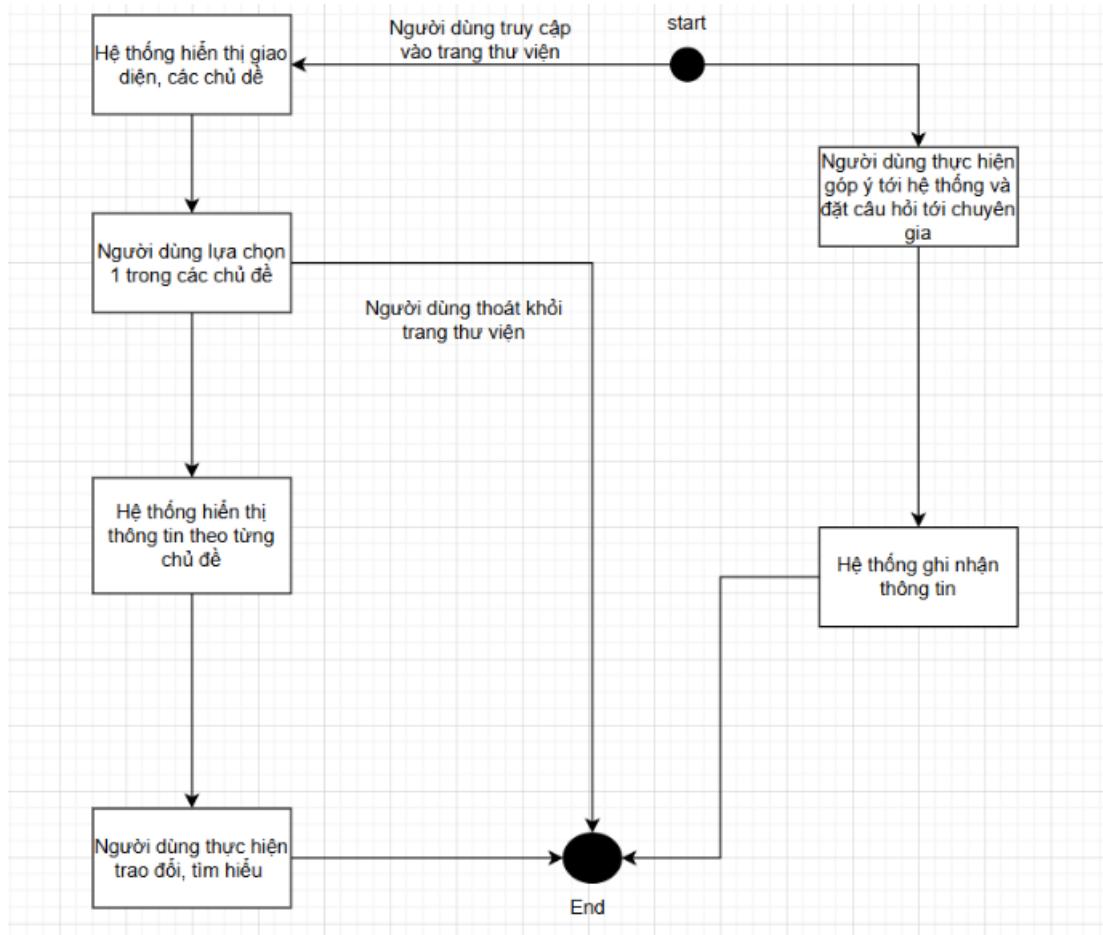
- Không.

##### Các luồng sự kiện:

- Luồng cơ bản:

- Người dùng truy cập vào trang bảng xếp hạng.
- Hệ thống hiển thị danh sách người dùng theo thứ hạng dựa trên điểm số.
- Người dùng xem vị trí của mình và danh sách người đứng đầu.

#### 4.2.9 Use Case 9: Cộng đồng



##### Các tác nhân

- Người dùng.
- Hệ thống.

##### Điều kiện kích hoạt ca sử dụng:

- Người dùng muốn trao đổi thông tin hoặc nhận hỗ trợ từ cộng đồng.

##### Tiền điều kiện:

- Người dùng đã đăng nhập vào hệ thống.

##### Hậu điều kiện:

- Bài viết, thông tin được cập nhật.

## Các luồng sự kiện:

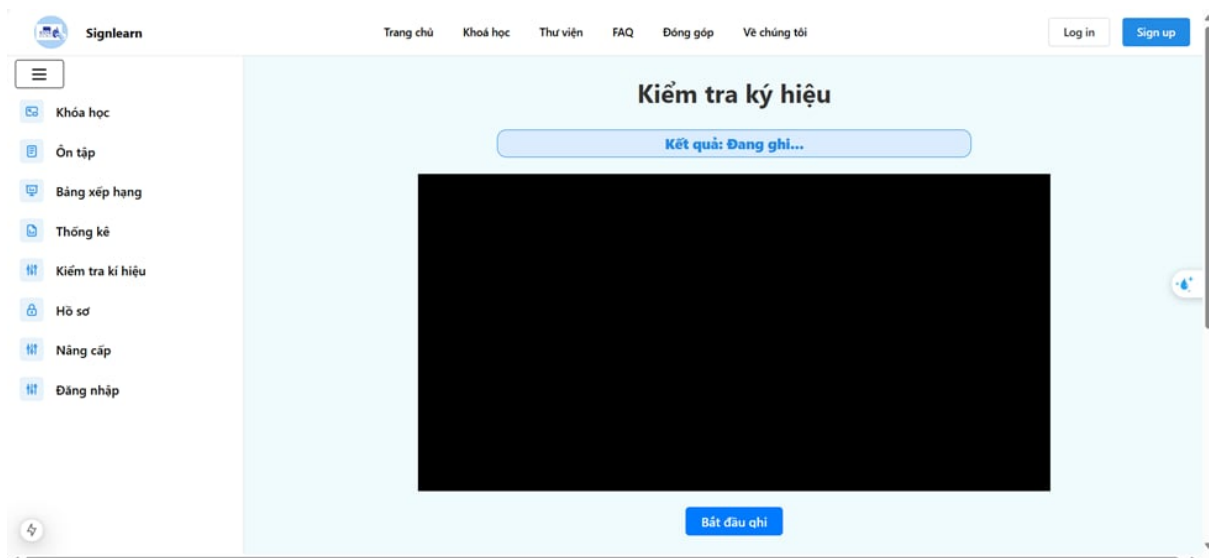
- Luồng cơ bản:
  - Người dùng truy cập vào trang thư viện.
  - Hệ thống hiển thị danh sách đề mục.
  - Người dùng chọn mục mình muốn tìm hiểu
  - Hệ thống xác nhận và trả về thông tin cần thiết
- Luồng mở rộng:
  - Người dùng phản hồi bài viết hoặc đặt câu hỏi tới hệ thống.

## 4.3 Góc nhìn các lớp tham gia

**Lớp giao diện (Client Layer):** Đảm bảo khả năng tiếp cận cho người dùng, thiết kế giao diện với biểu tượng dễ hiểu và hỗ trợ đa nền tảng.

### 4.3.1 Chức năng 1: Kiểm tra ký hiệu

Chức năng đầu tiên của ứng dụng là **Kiểm tra ký hiệu**. Với model CNN được xây dựng để phân tích cử chỉ ký hiệu thủ ngữ người dùng và trả về kết quả gần như là chính xác khi đã xử lý thông tin từ phí người dùng



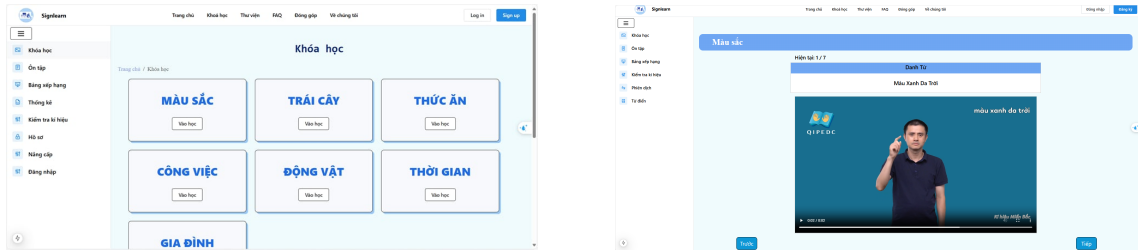
Hình 1: Ảnh chức năng kiểm tra tài liệu, camera ghi lại quá trình người dùng thực hiện ký hiệu thủ ngữ

Để thực hiện kiểm tra ký hiệu, người dùng cấp quyền sử dụng video cho ứng dụng, sau đó bấm nút ghi để bắt đầu quá trình. Sau đó người dùng thực hiện động tác cử chỉ

để quay lại và gửi tới hệ thống. Sau khi hệ thống phân tích sẽ trả về kết quả tương ứng với video gửi tới và gắn vào mục kết quả hiển thị ở trên màn hình.

### 4.3.2 Chức năng 2: Khoá học

Chức năng thứ hai của ứng dụng là **Khoá học**. Với hệ thống tài liệu phong phú được tổng hợp, hệ thống cho phép người dùng học tập từ vựng theo chủ đề, nhằm cải thiện khả năng ghi nhớ, liên kết từ ngữ và học tập một cách hiệu quả nhất.

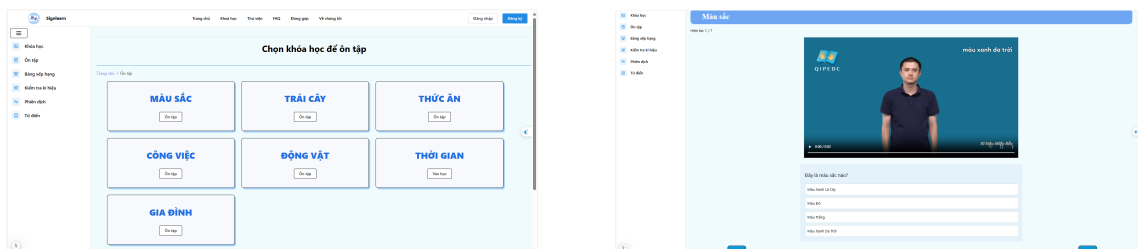


Hình 2: Ảnh chức năng khoá học, người dùng chọn chủ đề để thực hiện học tập

Để thực hiện học tập, người dùng tới trang khoá học và chọn chủ đề, Hệ thống sẽ trả về danh sách từ ngữ kèm theo video hướng dẫn cho từng chủ đề. Người dùng thực hiện theo để ghi nhớ và hiểu nội dung của từ ngữ.

### 4.3.3 Chức năng 3: Ôn tập

Chức năng thứ ba của ứng dụng là **ôn tập**. Tương tự với hệ thống khoá học, hệ thống ôn tập cho phép người dùng ôn tập lại các từ ngữ đã học cũng theo chủ đề đã có sẵn.



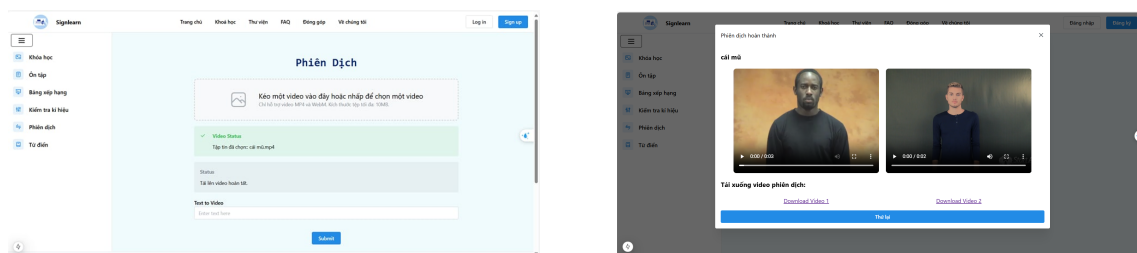
Hình 3: Ảnh chức năng ôn tập, người dùng chọn chủ đề để thực hiện học tập

Người dùng đến với trang ôn tập, chọn chủ đề và đợi hệ thống trả về danh sách video, với mỗi video, người dùng sẽ chọn nghĩa của từ đúng với cử chỉ trong video để kiểm tra lại trí nhớ của mình.

### 4.3.4 Chức năng 4: Phiên dịch

Chức năng thứ tư của ứng dụng là **Phiên dịch**. Mô hình dịch ngôn ngữ ký hiệu qua video không chỉ dừng lại ở việc nhận dạng mà còn thực hiện chuyển đổi ngôn ngữ ký

hiệu từ một ngôn ngữ nguồn (như ngôn ngữ ký hiệu tiếng Việt) sang một ngôn ngữ đích (ngôn ngữ tiếng anh).

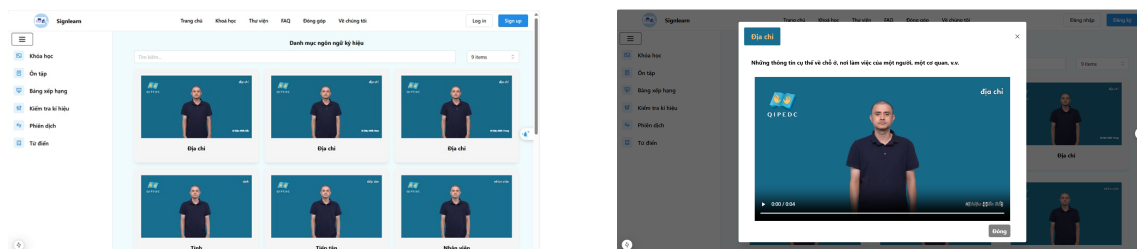


Hình 4: Ảnh chức năng khoá học, người dùng chọn chủ đề để thực hiện học tập

Người dùng đến với trang khoá học, chọn mục Phiên dịch để thực hiện, ở đây, người dùng có thể chọn kéo thả video để upload 1 video hướng dẫn thủ ngữ bằng tiếng việt, hệ thống sẽ phân tích và trả về video bằng tiếng anh và cho phép người dùng tải về để sử dụng và ôn tập.

#### 4.3.5 Chức năng 5: Từ điển

Chức năng tiếp theo của ứng dụng là **Từ điển**. Tận dụng nguồn tài liệu video phong phú về kí hiệu thủ ngữ, hệ thống xây dựng một trang từ điển hỗ trợ người dùng tra cứu từ vựng kèm theo video hướng dẫn và giải nghĩa ngữ cảnh sử dụng của từ



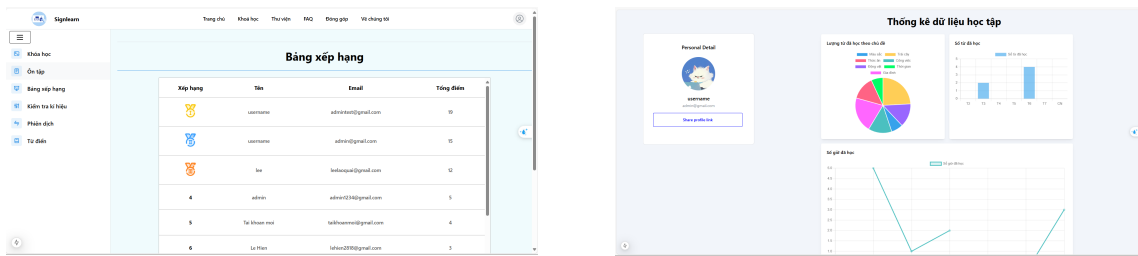
Hình 5: Ảnh chức năng Từ điển

Người dùng đến với trang khoá học, chọn mục Từ điển để thực hiện, ở đây, người dùng nhập từ vựng mình muốn tìm hiểu, nhấn vào thẻ, sẽ xuất hiện một khung chứa từ, nghĩa của từ kèm theo video hướng dẫn thực hiện kí hiệu thủ ngữ

#### 4.3.6 Chức năng 6: Thống kê và xếp hạng

Chức năng quan trọng của ứng dụng là **Thống kê và xếp hạng**. Để hỗ trợ người dùng đạt được hiệu quả tốt nhất trong việc học tập kí hiệu thủ ngữ, Hệ thống cho phép người dùng đăng nhập, thống kê quá trình học tập của người dùng cũng như xếp hạng để thúc đẩy tinh thần tự học của người dùng.





Hình 6: Ảnh chức năng Từ điển

### Lớp backend ứng dụng (Application Logic):

- Kiểm tra tính hợp lệ dữ liệu từ người dùng.
- Kết nối với backend để lưu trữ và truy xuất thông tin.

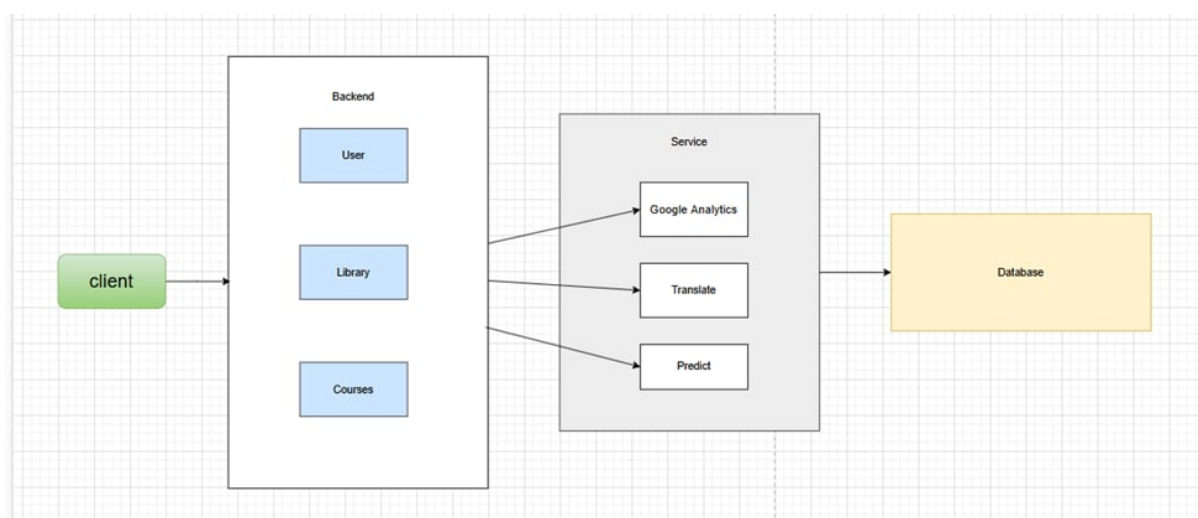
### Lớp xử lý dữ liệu (Data Layer):

- Lưu trữ dữ liệu người dùng, đăng ký, đăng nhập
- Cập nhật kết quả học tập vào cơ sở dữ liệu
- Lưu trữ và truy xuất dữ liệu học tập, cộng đồng

### Lớp dịch vụ (Services Layer):

- Phân tích video và nhận diện cử chỉ.
- Phân tích dữ liệu dữ liệu người dùng

Nhìn chung, luồng hoạt động ở phía backend sẽ như thế này



Hình 7: Minh họa luồng hoạt động của hệ thống và các services

## 4.4 Thiết kế cơ sở dữ liệu

Hệ thống sử dụng cơ sở dữ liệu quan hệ với các bảng chính sau:

### 4.4.1 Users (Người dùng):

- user\_id (Primary Key)
- name
- email
- password
- point

### 4.4.2 learningdata (Bài học):

- id (Primary Key)
- name
- target
- data
- id\_user (Foreign Key)
- status

### 4.4.3 revisedata (Kết quả thực hành):

- id (Primary Key)
- name
- target
- data
- id\_user (Foreign Key)
- status

#### **4.4.4 Libbase\_data (Cộng đồng):**

- id (Primary Key)
- title
- imgUrl
- content
- target

#### **4.4.5 Community\_data (Cộng đồng):**

- id (Primary Key)
- title
- imgUrl
- content
- contact
- id\_library (Foreign Key)

#### **4.4.6 document\_data (Cộng đồng):**

- id (Primary Key)
- title
- content
- id\_library (Foreign Key)

#### **4.4.7 Social\_data (Cộng đồng):**

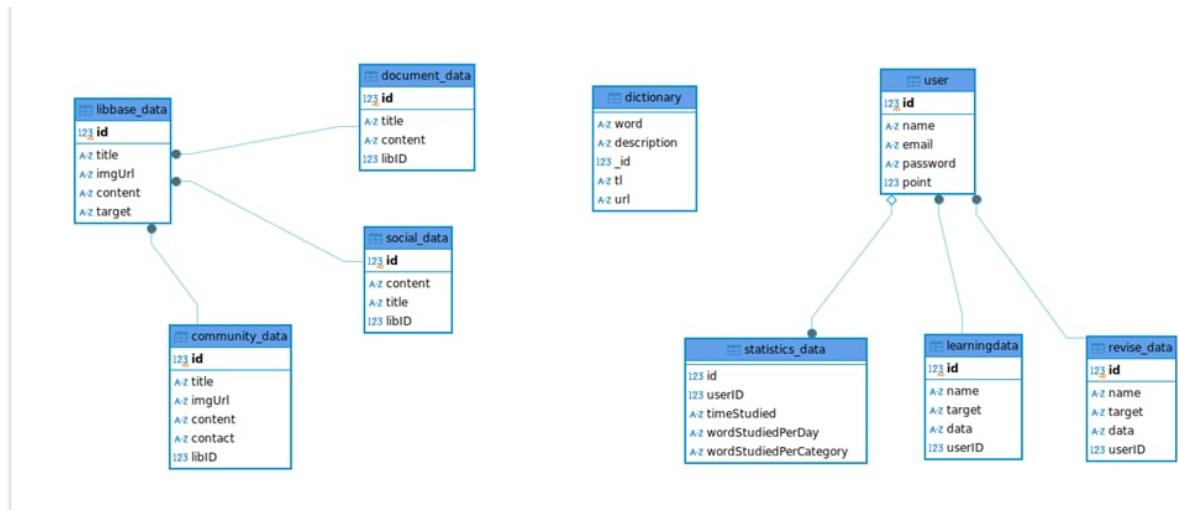
- id (Primary Key)
- title
- content
- id\_library (Foreign Key)

#### 4.4.8 Statistic\_data (Thống kê):

- id (Primary Key)
- timeStudied
- wordStudiedPerDay
- wordStudiedPerCategory
- userID (Foreign Key)

#### 4.4.9 Dictionary\_data (Từ điển):

- id (Primary Key)
- word
- descripton
- tl
- url



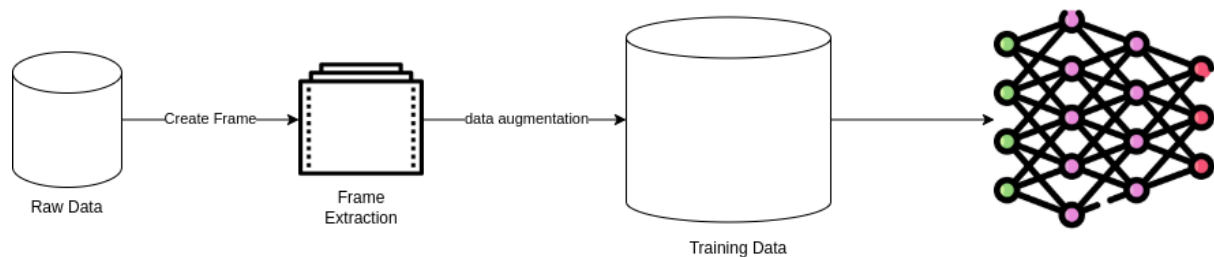
Hình 8: Minh hoạ cơ sở dữ liệu qua hệ

## 4.5 Thiết kế hệ thống thông minh

### 4.5.1 Hệ thống nhận diện ngôn ngữ ký hiệu

Hệ thống gồm mô hình nhận dạng ngôn ngữ ký hiệu qua video sử dụng kiến trúc "sequential", được thiết kế để phân tích chuỗi video và chuyển đổi chúng thành các đặc trưng nhằm phân loại ngôn ngữ ký hiệu. Đầu tiên, các đặc trưng không gian của từng

khung hình trong video được trích xuất thông qua tầng TimeDistributed, chuyển dữ liệu từ dạng không gian (7, 7, 1280) sang dạng vector (1280) đại diện cho mỗi khung hình. Sau đó, hai tầng GRU (Gated Recurrent Unit) được áp dụng để xử lý chuỗi đặc trưng, giúp mã hóa thông tin tuần tự giữa các khung hình và hiểu được ngữ cảnh toàn bộ video. Tiếp theo, một tầng Batch Normalization được sử dụng để chuẩn hóa dữ liệu, tăng tốc độ hội tụ và ổn định quá trình huấn luyện. Đầu ra của tầng GRU cuối cùng là một vector đại diện cho toàn bộ video, được đưa qua một chuỗi các tầng Dense với số lượng đơn vị giảm dần từ 64, 32 đến 20, nhằm học các đặc trưng quan trọng và phân loại ngôn ngữ ký hiệu. Để tránh hiện tượng overfitting, các tầng Dropout được chèn giữa các tầng Dense. Với tổng số 3.854.132 tham số, mô hình này có thể áp dụng hiệu quả trong các hệ thống hỗ trợ giao tiếp cho người khiếm thính, chẳng hạn như dịch ngôn ngữ ký hiệu thành văn bản hoặc giọng nói, đồng thời mở ra tiềm năng tích hợp các kỹ thuật tiên tiến như Attention để tăng cường hiệu suất nhận dạng.



Hình 9: Đào tạo mô hình nhận diện ngôn ngữ kí hiệu

Model: "sequential"		
Layer (type)	Output Shape	Param #
time_distributed (TimeDistri	(None, 10, 7, 7, 1280)	2257984
time_distributed_1 (TimeDist	(None, 10, 1280)	0
gru (GRU)	(None, 10, 256)	1181184
batch_normalization (BatchNo	(None, 10, 256)	1024
gru_1 (GRU)	(None, 256)	394752
dense (Dense)	(None, 64)	16448
dropout (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 32)	2080
dropout_1 (Dropout)	(None, 32)	0
dense_2 (Dense)	(None, 4)	132
Total params: 3,853,604		
Trainable params: 3,818,980		
Non-trainable params: 34,624		

Hình 10: Kiến trúc mô hình nhận diện ngôn ngữ kí hiệu

#### 4.5.2 Hệ thống dịch ngôn ngữ kí hiệu

Mô hình dịch ngôn ngữ ký hiệu qua video không chỉ dừng lại ở việc nhận dạng mà còn thực hiện chuyển đổi ngôn ngữ ký hiệu từ một ngôn ngữ nguồn (như ASL - American Sign Language) sang một ngôn ngữ đích (như ngôn ngữ ký hiệu tiếng Việt). Hệ thống này được chia thành ba giai đoạn chính:

**Nhận dạng từ ngôn ngữ ký hiệu qua video** Hệ thống nhận đầu vào là một video biểu diễn ngôn ngữ ký hiệu. Đầu tiên, mô hình nhận dạng ngôn ngữ ký hiệu cấp độ từ vựng được áp dụng để phân tích video và dự đoán các từ hoặc cụm từ trong ngôn ngữ ký hiệu nguồn. Ví dụ:

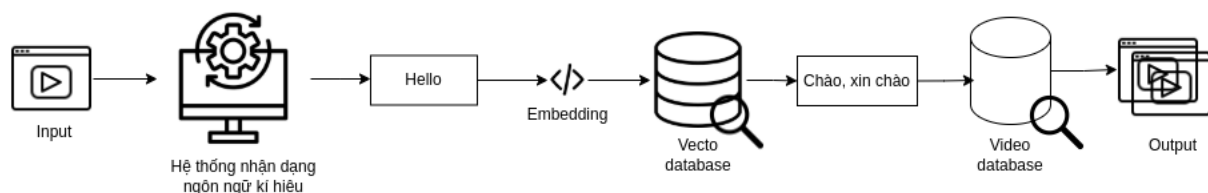
- **Video đầu vào:** Một người đang ký hiệu từ "hello" trong ASL.
- **Đầu ra từ giai đoạn 1:** Chuỗi dự đoán ký hiệu dạng text (hello).

**Dịch từ ngôn ngữ nguồn sang ngôn ngữ đích** Hệ thống nhận đầu vào là một video biểu diễn ngôn ngữ ký hiệu. Đầu tiên, mô hình nhận dạng ngôn ngữ ký hiệu cấp độ từ

vững được áp dụng để phân tích video và dự đoán các từ hoặc cụm từ trong ngôn ngữ ký hiệu nguồn. Ví dụ:

- **Mô hình dịch:** Mô hình dịch được xây dựng dựa trên kiến trúc Seq2Seq kết hợp với Attention, cho phép học được ngữ cảnh giữa các từ và dịch chính xác theo ngữ nghĩa.
- **Embedding:** Từ hoặc cụm từ dự đoán được ánh xạ vào không gian embedding. Trong không gian này, mô hình tìm kiếm các từ có ngữ nghĩa gần nhất trong cơ sở dữ liệu ngôn ngữ ký hiệu đích để đảm bảo tính chính xác và tự nhiên.
- **Input** (ngôn ngữ nguồn): hello
- **Output** (ngôn ngữ đích): xin chào, chào, hi.

Sau đó hệ thống sẽ tìm kiếm video có title giống với Output khi dịch.



Hình 11: Đào tạo mô hình nhận diện ngôn ngữ kí hiệu

## 5 Phát triển và Công nghệ sử dụng

### 5.1 Công nghệ và công cụ phát triển

- **Nền tảng Frontend website:** Framework Nextjs, thư viện UI Mantine để phát triển ứng dụng cho website
- **Backend:** Framework web python Flask
- **Phân tích dữ liệu:** Google Analytics tích hợp để theo dõi hành vi người dùng.
- **Nền tảng xử lý video và nhận diện cử chỉ:** Hệ thống sử dụng mô hình Convolutional Neural Network (CNN) kết hợp với MobileNetV2 và các lớp xử lý video như TimeDistributed và GRU để trích xuất, phân loại cử chỉ từ video. Mô hình này giúp nhận diện và phân tích các động tác của người dùng một cách hiệu quả.
- **Tăng cường dữ liệu (Data Augmentation):** Thực hiện các kỹ thuật tăng cường dữ liệu để mở rộng tập dữ liệu, cải thiện tính đa dạng của mẫu cử chỉ, giúp mô hình học tốt hơn và giảm thiểu vấn đề overfitting.

## 5.2 Quy trình phát triển

- **Giai đoạn 1:** Phân tích yêu cầu và thiết kế hệ thống.
- **Giai đoạn 2:** Phát triển frontend và backend.
- **Giai đoạn 3:** Tích hợp các tính năng và thiết kế giao diện.
- **Giai đoạn 4:** Kiểm thử và sửa lỗi.
- **Giai đoạn 5:** Triển khai và bảo trì ứng dụng.

## 5.3 Các thách thức kỹ thuật

- **Xử lý cử chỉ ký hiệu:** Đảm bảo nhận diện chính xác các cử chỉ phức tạp, mô hình học sâu xây dựng phải được training với tập dữ liệu đầy đủ, đảm bảo trả về kết quả đúng nhất cho người dùng.
- **Tối ưu hóa hiệu suất:** Đảm bảo ứng dụng chạy mượt mà trên nhiều thiết bị khác nhau, không gặp vấn đề trong những trường hợp đặc biệt như giờ cao điểm, gây ảnh hưởng đến trải nghiệm người dùng.
- **Bảo mật dữ liệu:** Bảo vệ thông tin cá nhân và dữ liệu học tập của người dùng.
- **Thiếu hụt dữ liệu về ngôn ngữ ký hiệu tại Việt Nam:** Hiện tại, nguồn dữ liệu ký hiệu tại Việt Nam còn hạn chế, gây khó khăn trong quá trình huấn luyện mô hình. Cần áp dụng các phương pháp thu thập, tăng cường dữ liệu và cộng tác với các chuyên gia để cải thiện độ phong phú của tập dữ liệu.

## 5.4 Phương pháp kiểm thử ứng dụng

- **Kiểm thử chức năng:** Đảm bảo tất cả các tính năng hoạt động đúng như thiết kế.
- **Kiểm thử tương thích:** Đảm bảo ứng dụng hoạt động tốt trên các thiết bị và hệ điều hành khác nhau.
- **Kiểm thử người dùng:** Thu thập phản hồi từ một nhóm người dùng thử để cải thiện trải nghiệm.
- **Kiểm thử bảo mật:** Đánh giá và bảo vệ hệ thống khỏi các mối đe dọa bảo mật.



## 6 Kết quả và Đánh giá

### 6.1 Đánh giá hiệu quả của ứng dụng

Ứng dụng SignLearning đã đạt được mục tiêu cung cấp một nền tảng học ngôn ngữ ký hiệu hiệu quả, giúp người dùng tăng cường kỹ năng giao tiếp và tự tin hơn trong việc sử dụng ngôn ngữ ký hiệu.

### 6.2 Các điểm cải tiến và tiềm năng mở rộng

- **Cải tiến chất lượng nội dung:** Thêm nhiều video hướng dẫn chất lượng cao và cập nhật thường xuyên.
- **Tích hợp trí tuệ nhân tạo:** Sử dụng AI để cung cấp phản hồi tự động và nhận diện cử chỉ chuẩn xác hơn.
- **Mở rộng ngôn ngữ ký hiệu:** Phát triển các phiên bản cho các hệ ngôn ngữ ký hiệu khác nhau.

## 7 Thành Tựu

### 7.1 Tóm tắt thành tựu đạt được

Ứng dụng đã được triển khai thành công và có thể thực hiện các chức năng chính như Học và Ôn tập, cùng với phiên dịch, từ điển và đặc biệt là Kiểm tra ký hiệu. Các thử nghiệm cho thấy ứng dụng có khả năng phục vụ nhu cầu của người dùng khiếm thính một cách hiệu quả. Tuy nhiên, để nâng cao trải nghiệm người dùng, các chức năng này vẫn cần được tối ưu hóa về mặt tốc độ và độ chính xác, số lượng từ ngữ phong phú và đầy đủ hơn.

### 7.2 Hướng phát triển trong tương lai

Trong tương lai, sau khi đã hoàn thành kiểm thử chức năng cơ bản, nhóm đưa sản phẩm ra thử nghiệm tại một phần nhỏ của cộng đồng, nhận kết quả và mong muốn hoàn thành hơn nữa, phát triển thêm nhằm đáp ứng hiệu quả tốt nhất cho người dùng. Về phương hướng phát triển, nhóm dự định tiếp tục nâng cao chất lượng ứng dụng bằng cách tích hợp các công nghệ mới như trí tuệ nhân tạo, mở rộng nội dung học tập và phát triển thêm các tính năng hỗ trợ người dùng, nhằm tạo ra một môi trường học tập toàn diện và tương tác hơn.

## 8 Bảng phân chia công việc

Họ và tên	MSSV	Đóng góp	Công việc
Lê Thế Hiền (Nhóm trưởng)	22028101	25%	Xây dựng, triển khai model phân tích phiên dịch cử chỉ kí hiệu thủ ngữ. Hỗ trợ thiết kế API Endpoint backend Làm báo cáo, thuyết trình, quay video demo sản phẩm.
Lê Văn Đức	22028041	25%	Tìm hiểu, thiết kế cài đặt các chức năng liên quan tới chức năng phân tích, phiên dịch, từ điển. Hỗ trợ thiết kế, xây dựng các chức năng khác của dự án Thiết kế API Endpoing backend Viết báo cáo, hỗ trợ quay video demo
Nguyễn Quang Khải	22028299	25%	Tìm kiếm các công nghệ liên quan chức năng liên quan tới khoá học, ôn tập, thống kê tài liệu người dùng Hỗ trợ xây dựng cài đặt các chức năng liên qua tới tài liệu FAQ, about Thiết kế slide
Phạm Quý Sơn	22028296	25%	Tìm kiếm cài đặt các công nghệ liên quan trang chủ, tài liệu, trang FAQ, about Thực hiện cài đặt hỗ trợ cài đặt chức năng ôn tập, khoá học. Thiết kế slide, hỗ trợ quay video demo

Bảng 1: Phân chia công việc trong nhóm

Các thành viên trong nhóm cũng đã cùng nhau phối hợp để giải quyết những vấn đề phát sinh trong quá trình triển khai, đảm bảo tiến độ và chất lượng của dự án.

## 9 Phụ lục và Tài liệu tham khảo

### 9.1 Phụ lục

- **Hình ảnh giao diện người dùng:** Bao gồm các wireframes và mẫu thiết kế cuối cùng.

- **Sơ đồ hệ thống:** Mô tả kiến trúc backend và frontend của ứng dụng.
- **Bảng biểu tiến độ phát triển:** Ghi lại các mốc quan trọng trong quá trình phát triển dự án.

## 9.2 Tài liệu tham khảo

- Smith, J., & Doe, A. (2020). Human-Computer Interaction: An Empirical Research Perspective. Springer.
- Nguyen, T. H., & Le, P. M. (2019). Applied Sign Language Technologies. IEEE Transactions on Affective Computing.
- Website học tập ngôn ngữ ký hiệu QIPEDC do Quỹ Hợp tác toàn cầu tài trợ theo kết quả đầu ra (GPRBA) ủy thác: <https://qipcdc.moet.gov.vn/dictionary>
- Các tài liệu tham khảo về ngôn ngữ ký hiệu Việt Nam từ các nguồn uy tín như Viện Điện tử và Tin học Việt Nam.
- Li, Dongxu and Rodriguez, Cristian and Yu, Xin and Li, Hongdong. Word-level Deep Sign Language Recognition from Video: A New Large-scale Dataset and Methods Comparison
- Li, Dongxu and Yu, Xin and Xu, Chenchen and Petersson, Lars and Li, Hongdong. Li, Dongxu and Yu, Xin and Xu, Chenchen and Petersson, Lars and Li, Hongdong
- Li, Dongxu and Xu, Chenchen and Yu, Xin and Zhang, Kaihao and Swift, Benjamin and Suominen, Hanna and Li, Hongdong. Tspnet: Hierarchical feature learning via temporal semantic pyramid for sign language translation
- Li, Dongxu and Xu, Chenchen and Liu, Liu and Zhong, Yiran and Wang, Rong and Petersson, Lars and Li, Hongdong. Transcribing natural languages for the deaf via neural editing programs
- G Adithya V., Vinod P. R., Usha Gopalakrishnan, “Artificial Neural Network Based Method for Indian Sign Language Recognition”, IEEE Conference on Information and Communication Technologies (ICT), 2013, pp. 1080-1085.
- Bach Duy Khuat, Duong Thai Phung, Ha Thi Thu Pham, Anh Ngoc Bui, Son Tung Ngo. Authors Info, Claims. Vietnamese sign language detection using Mediapipe
- The Duy Bui, Long Thang Nguyen. Recognizing Postures in Vietnamese Sign Language With MEMS Accelerometers
- AH Vo, VH Pham, BT Nguyen. Deep learning for Vietnamese Sign Language recognition in video sequence