

TRƯỜNG ĐẠI HỌC BÁCH KHOA  
KHOA CÔNG NGHỆ THÔNG TIN



## ĐỒ ÁN TỐT NGHIỆP

*Tên đề tài:*

**XÂY DỰNG WEBSITE TỪ ĐIỂN VIỆT NHẬT CÓ  
ỨNG DỤNG MÔ HÌNH AI DỊCH THUẬT**

SVTH : Hoàng Nguyễn Nhật Minh. 21TCLC\_NHAT2  
GVHD: ThS. Nguyễn Thế Xuân Ly

*Đà Nẵng, 2025*

## NHẬN XÉT ĐỒ ÁN TỐT NGHIỆP

### 1. Thông tin chung:

1. Họ và tên sinh viên: Hoàng Nguyễn Nhật Minh  
2. Lớp: 21TCLC\_NHAT2 Số thẻ SV: 102210361  
3. Tên đề tài: Xây dựng website từ điển Việt Nhật có ứng dụng mô hình AI dịch thuật.  
4. Người hướng dẫn: ThS Nguyễn Thế Xuân Ly Học hàm/ học vị: Thạc sĩ

### II. Nhận xét, đánh giá đồ án tốt nghiệp:

1. Về tính cấp thiết, tính mới, khả năng ứng dụng của đề tài: (điểm tối đa là 2đ)  
.....  
.....  
2. Về kết quả giải quyết các nội dung nhiệm vụ yêu cầu của đồ án: (điểm tối đa là 4đ)  
.....  
.....  
3. Về hình thức, cấu trúc, bố cục của đồ án tốt nghiệp: (điểm tối đa là 2đ)  
.....  
.....  
4. Đề tài có giá trị khoa học/ có bài báo/ giải quyết vấn đề đặt ra của doanh nghiệp hoặc nhà trường: (điểm tối đa là 1đ)  
.....  
.....  
5. Các tồn tại, thiếu sót cần bổ sung, chỉnh sửa:  
.....  
.....

### III. Tinh thần, thái độ làm việc của sinh viên: (điểm tối đa 1đ)

.....

### IV. Đánh giá:

1. Điểm đánh giá: ...../10 (lấy đến 1 số lẻ thập phân)  
2. Đề nghị: ☐ Được bảo vệ đồ án ☐ Bổ sung để bảo vệ ☐ Không được bảo vệ

Đà Nẵng, ngày      tháng      năm 2025

**Người hướng dẫn**

# BẢN TÓM TẮT BẰNG TIẾNG NHẬT

日本語-英語辞書サイト (AI 翻訳モデル搭載)

Japanese-English dictionary website with AI-based translation model

名前: HOANG NGUYEN NHAT MINH (ホアン・グエン・ニャット・ミン)

学生番号: 102210361

指導教官: NGUYEN THE XUAN LY

所属: ダナン大学・工科大学・情報学部

**Abstract:** In modern times, with the increasing number of Japanese language learners, improving the efficiency of vocabulary lookup and comprehension of original texts has become an urgent issue. This app integrates AI-powered natural full-text translation with OCR technology, enabling users to instantly obtain accurate translations simply by capturing images of Japanese texts from books or materials. Furthermore, it features the scientifically proven FSRs (Forgetting Spaced Repetition System) to optimize memory retention. Through personalized word card creation and management, automatic review timing notifications, and a progress dashboard, the app supports active and long-term vocabulary acquisition.

## 1. はじめに

近年、日本語学習者の増加に伴い、効率的かつ能動的に語彙を習得できるツールへのニーズが高まっています。本プロジェクトでは、ベトナム語⇄日本語の双方向辞書機能に加え、AI 翻訳エンジンと OCR を統合したウェブサイトを開発します。さらに、FSRS（忘却間隔反復システム）による復習管理とカスタマイズ可能な単語カードで、学習者が無理なく長期記憶へ定着できる環境を提供します。

## 2. システム分析、設計

### 2.1. 機能分析

管理者:

- 語彙および漢字の追加・編集・削除
- ユーザー一覧およびユーザー詳細の閲覧
- ユーザー数、語彙数、漢字数などのデ

ータをグラフで表示

ユーザー:

- アカウント情報の管理
- 単語カードおよび復習カードの作成・管理
- 語彙と漢字の学習機能
- 学習進捗や統計データの閲覧
- 語彙・漢字の検索
- 検索履歴の閲覧
- 日本語⇄ベトナム語の短文または画像翻訳

ゲスト:

- 語彙・漢字の検索
- 検索履歴の閲覧
- 日本語⇄ベトナム語の短文または画像翻訳

### 2.2. ユースケース図

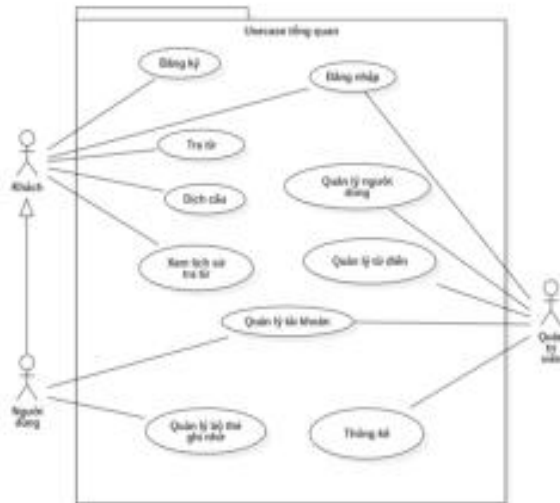


図 1 概要のユースケース図

## 2.3. データベース図



図 2 データベース図

## 3. 実験と結果

### 3.1. 実験環境

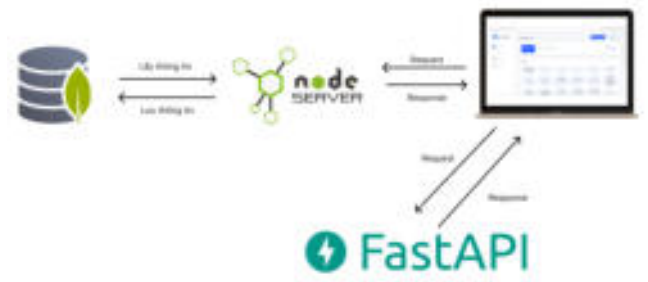


図 3 システム概要

### 3.2. 結果

#### 語彙検索画面

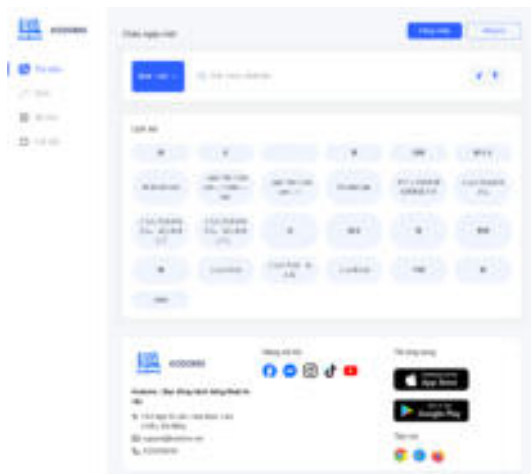


図 4 語彙検索画面

#### 文翻訳画面



図 5 文翻訳画面

## カードセット画面

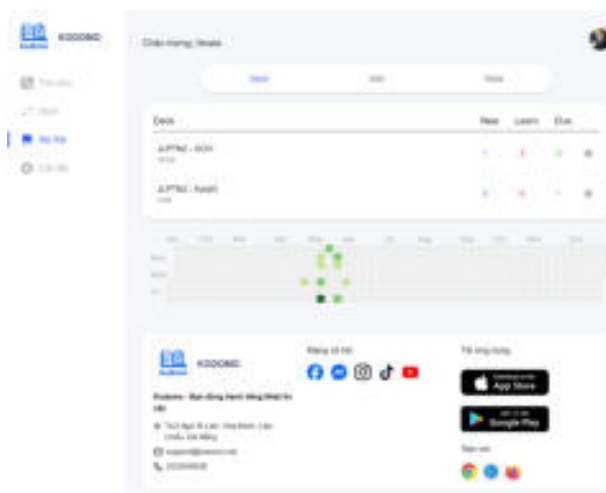


図6 カードセット画面

## カード学習画面



図7 カード学習画面

## 4. 結論と開発の方向性

### 4.1. 達成結果

プロジェクトの理論学習および開発実施を通じて、当初設定した目標を一定程度達成しました。

#### 理論について:

- データベースの分析・設計・構築スキルを習得（要件定義、データ構造設計、

DB 展開を含む）。

- TypeScript、Python、および ExpressJS、ReactJS、FastAPI といったフレームワーク／ライブラリを理解し、システム開発に応用。
- 現行の NLP モデル（Transformer 系）による言語処理原理を学び、翻訳機能への応用を試行。

#### 申請に関して:

- ウェブサイトは当初提案した全機能を実装。
- ユーザーは迅速に語彙・漢字を検索でき、AI 翻訳機能により文脈に即した自然な訳文を取得可能。
- FRSR（忘却間隔反復システム）を用いたフラッシュカード学習により、効率的な語彙記憶が実現。

### 4.2. 欠点

- 管理者機能：時間的制約により、一部機能が未実装。
- 性能：同時アクセス増加時にレスポンスが低下する可能性。
- UI／UX：画面デザインが簡素で、視認性・操作性の向上が必要。
- 辞書コンテンツ：ベトナム語-日本語データが不十分で、一部語句が欠落。
- フラッシュカード：現在は基本情報（意味・読み・例文）のみ登録可能で、ユーザー任意の項目追加ができない。
- AI 翻訳モデル：訓練データ量・リソース不足により長文翻訳や文体表現で精度が課題。

### 4.3. 開発の方向性

- 辞書データベースの拡張：ベトナム語-日本語のバイリンガルデータを継続収集・クレンジングし、項目数・品質を向上。学習者コミュニティと協働して

実用的な用例・語彙を補完。

- **Transformer** モデルの改善：ハードウェア強化およびクラウド環境で大規模データを用いた再訓練を実施。文長多様な翻訳コーパスを拡充し、長文処理性能を向上。
- フラッシュカード学習機能の多様化：画像カード、コンテキスト例カード、音声テストカードなど新形式を統合。学習履歴・ユーザー嗜好に基づくカード自動提案機能を開発。
- 画像から文字抽出機能の強化：先進 OCR モデル導入とアルゴリズム最適化で認識精度を向上。手書き・印刷文字

を含む多様な画像データセットを収集し学習。

- モバイルアプリの開発：フラッシュカード学習や語彙検索を手軽に行える専用アプリを提供。
- 評価・フィードバックシステムの構築：辞書エントリへのコメント・評価機能を実装。ユーザー提案をもとに語彙推薦・コンテンツ更新を行うモデレーション体制を整備。
- 新技術の研究・統合：GPT、BERT など最新 NLP モデルの適用可能性を検討。学習支援チャットボットやテーマ別テスト機能の実装を目指す。

## TÓM TẮT

Tên đề tài: Xây dựng website từ điển Việt – Nhật có ứng dụng mô hình AI dịch thuật

Sinh viên thực hiện: Hoàng Nguyễn Nhật Minh

Số thẻ SV: 102210361

Lớp: 21TCLC\_NHAT2

Tóm tắt đề tài:

Trong bối cảnh hiện nay, khi nhu cầu học tiếng Nhật ngày càng gia tăng và việc tra cứu từ vựng, hiểu nhanh văn bản gốc trở thành một thách thức đối với nhiều người học, việc sở hữu một công cụ hỗ trợ toàn diện là vô cùng cần thiết. Tuy nhiên, hầu hết các ứng dụng từ điển hiện có vẫn chỉ dừng lại ở việc tra từ một cách thụ động, chưa kết hợp sâu các phương pháp học tập khoa học hay công nghệ thông minh để nâng cao hiệu quả tiếp thu. Vì vậy, việc phát triển một ứng dụng từ điển tiếng Nhật thông minh, tích hợp AI và phương pháp ghi nhớ ngắt quãng sẽ giải quyết đúng nhu cầu thực tế, giúp người dùng không chỉ tra cứu nhanh mà còn ghi nhớ lâu dài, học tập chủ động hơn.

Ứng dụng này sẽ cung cấp nền tảng tra cứu từ vựng, áp dụng AI giúp dịch nguyên câu theo ngữ nghĩa tự nhiên, khắc phục tình trạng hiểu sai khi dịch từng từ rời rạc. Người học không còn lo ngại về việc dịch máy cứng nhắc; thay vào đó, bản dịch sẽ gần gũi và chính xác hơn, hỗ trợ họ tiếp cận nội dung gốc một cách trực quan. Ngoài ra, bằng việc tích hợp công nghệ OCR (Optical Character Recognition), người dùng có thể chụp ảnh văn bản tiếng Nhật từ sách, hay tài liệu bất kỳ, và ngay lập tức nhận được bản dịch, giúp việc học và tra cứu trở nên tiện lợi tối đa.

Để đảm bảo người dùng ghi nhớ từ vựng hiệu quả, ứng dụng áp dụng hệ thống “Forgotten Spaced Repetition System” (FSRS) – một phương pháp ghi nhớ ngắt quãng đã được chứng minh về mặt khoa học giúp đưa từ vựng vào bộ nhớ dài hạn. Hệ thống này sẽ tự động tính toán khoảng cách ôn tập dựa trên tần suất và kết quả luyện tập, nhắc nhở người học ôn lại đúng thời điểm để giảm thiểu tình trạng “học trước quên sau”. Song song đó, người dùng có thể tự tạo và quản lý bộ thẻ từ vựng cá nhân hóa, đánh dấu những từ khó, theo dõi tiến độ học tập và nhận gợi ý về danh sách từ cần ôn tập tiếp theo dựa trên mức độ ghi nhớ và mục tiêu cá nhân.

### NHIỆM VỤ ĐỒ ÁN TỐT NGHIỆP

Họ tên sinh viên: Hoàng Nguyễn Nhật Minh

Số thẻ sinh viên: 102210361

Lớp: 21TCLC\_NHAT2

Khoa: Công nghệ thông tin

Ngành: CNTT Việt-Nhật

1. Tên đề tài đồ án:

Xây dựng website từ điển Việt Nhật có ứng dụng mô hình AI dịch thuật

2. Đề tài thuộc diện: ☐ Có ký kết thỏa thuận sở hữu trí tuệ đối với kết quả thực hiện

3. Các số liệu và dữ liệu ban đầu:

Không có

4. Nội dung các phần thuyết minh và tính toán:

Nội dung các phần thuyết minh bao gồm:

- Tổng quan đề tài: Giới thiệu tổng quan, mục đích và ý nghĩa của đề tài, công nghệ, kỹ thuật, công cụ được sử dụng và những kết quả dự kiến đạt được.
- Cơ sở lý thuyết: Trình bày những cơ sở lý thuyết được áp dụng trong đề tài.
- Phân tích thiết kế hệ thống: Phân tích các yêu cầu về chức năng và phi chức năng của hệ thống và triển khai thiết kế hệ thống thông qua các biểu đồ thiết kế hệ thống.
- Triển khai mô hình và thuật toán: Trình bày các bước triển khai mô hình và thuật toán trong ứng dụng.
- Triển khai thực tế: Trình bày môi trường triển khai và kết quả thực tế đạt được.
- Kết luận và hướng phát triển: Đánh giá kết quả đạt được, chưa đạt được và đưa ra những định hướng phát triển thêm trong tương lai.
- Tài liệu tham khảo: Liệt kê các tài liệu tham khảo sử dụng trong đề tài.

5. Các bản vẽ, đồ thị (ghi rõ các loại và kích thước bản vẽ):

Không có

6. Họ tên người hướng dẫn: ThS Nguyễn Thế Xuân Ly

7. Ngày giao nhiệm vụ đồ án:

8. Ngày hoàn thành đồ án:

Đà Nẵng, ngày      tháng      năm 2025.

Trưởng Bộ môn .....

Người hướng dẫn

GVHD: ThS Nguyễn Thế Xuân Ly

## LỜI NÓI ĐẦU

Đề tài “Xây dựng website từ điển Việt Nhật có ứng dụng mô hình AI dịch thuật” là kết quả của quá trình nỗ lực và tâm huyết của em trong thời gian theo học tại Khoa Công nghệ Thông tin – Trường Đại học Bách khoa – Đại học Đà Nẵng. Đây cũng là đề tài đồ án tốt nghiệp của em, đánh dấu một bước quan trọng trong quá trình hoàn thiện chương trình đại học và tích lũy kinh nghiệm thực tiễn. Thông qua việc nghiên cứu và phát triển ứng dụng này, em mong muốn đóng góp phần nhỏ bé của mình vào việc tạo ra công cụ hỗ trợ học tiếng Nhật hiệu quả hơn, đồng thời củng cố và phát triển những kiến thức, kỹ năng đã được học tập.

Em xin gửi lời tri ân sâu sắc đến Nhà trường và quý thầy cô trong Khoa Công nghệ Thông tin, những người đã truyền đạt cho em những kiến thức nền tảng quý báu về lập trình, trí tuệ nhân tạo và thiết kế giao diện người dùng. Những bài giảng, buổi hướng dẫn thực hành và các buổi trao đổi chuyên môn với thầy cô đã giúp em định hình tư duy, phương pháp nghiên cứu, cũng như cung cấp cho em các công cụ cần thiết để thực hiện đề tài. Đặc biệt, em xin được bày tỏ lòng biết ơn chân thành đến ThS. Nguyễn Thế Xuân Ly, người đã tận tâm hướng dẫn, hỗ trợ và góp ý từng bước trong quá trình em phát triển ứng dụng.

Trong quá trình thực hiện đề tài, chắc chắn sẽ không tránh khỏi những hạn chế và thiếu sót nhất định. Em rất mong nhận được những góp ý, phản hồi từ thầy cô và bạn bè để có thể hoàn thiện hơn về mặt tính năng, giao diện và thuật toán. Em cũng hy vọng ứng dụng này sẽ thực sự hữu ích cho cộng đồng người học tiếng Nhật, góp phần nâng cao hiệu quả ghi nhớ và năng lực ngôn ngữ. Cuối cùng, em xin kính chúc quý thầy cô, anh chị và các bạn sức khỏe, hạnh phúc và ngày càng gặt hái nhiều thành công trong sự nghiệp và cuộc sống. Trân trọng!

Sinh viên thực hiện

Hoàng Nguyễn Nhật Minh

## LỜI CAM ĐOAN

### **Em xin cam đoan:**

1. Nội dung trong đồ án này em thực hiện dưới sự hướng dẫn trực tiếp của ThS Nguyễn Thế Xuân Ly. Các kết quả trong đồ án đều được thực hiện trong quá trình em đang học tập tại Khoa Công Nghệ Thông Tin, trường Đại học Bách Khoa Đà Nẵng.
2. Các tham khảo dùng trong đồ án đều được trích dẫn rõ ràng tên tác giả, tên công trình, thời gian, địa điểm công bố, được kèm với đường dẫn hợp lệ.
3. Nếu có những sao chép không hợp lệ, vi phạm quy chế đào tạo, em xin chịu hoàn toàn trách nhiệm.

*Đà Nẵng, ngày.....tháng.....năm 2025*

**Sinh viên thực hiện**

Hoàng Nguyễn Nhật Minh

## MỤC LỤC

|                                                                         |           |
|-------------------------------------------------------------------------|-----------|
| BẢN TÓM TẮT BẰNG TIẾNG NHẬT .....                                       | i         |
| TÓM TẮT .....                                                           | v         |
| NHIỆM VỤ ĐỒ ÁN TỐT NGHIỆP .....                                         | vi        |
| LỜI NÓI ĐẦU .....                                                       | vii       |
| LỜI CAM ĐOAN .....                                                      | viii      |
| MỤC LỤC.....                                                            | ix        |
| DANH MỤC HÌNH ẢNH .....                                                 | xii       |
| DANH MỤC BẢNG BIỂU .....                                                | xiv       |
| <b>MỞ ĐẦU .....</b>                                                     | <b>1</b>  |
| <b>CHƯƠNG 1. CƠ SỞ LÝ THUYẾT .....</b>                                  | <b>5</b>  |
| 1.1. Tổng quan về mô hình Transformer .....                             | 5         |
| 1.1.1. Vấn đề của mô hình tuần tự (RNN, LSTM) .....                     | 5         |
| 1.1.2. Cách Transformer xử lý toàn cục bằng Self-Attention.....         | 5         |
| 1.1.3. Kiến trúc tổng thể .....                                         | 5         |
| 1.1.4. Cơ chế Attention và Multi-head Attention.....                    | 8         |
| 1.1.5. Triển khai và ứng dụng .....                                     | 9         |
| 1.1.6. Ưu điểm .....                                                    | 9         |
| 1.1.7. Hạn chế .....                                                    | 9         |
| 1.2. Tổng quan về thuật toán FSRS.....                                  | 9         |
| 1.2.1. Đường cong quên và sử phảng dần qua mỗi lần lặp.....             | 9         |
| 1.2.2. Nguyên lý hoạt động của FSRS và áp dụng vào đường cong quên. ... | 11        |
| <b>CHƯƠNG 2. CÔNG CỤ PHÁT TRIỂN.....</b>                                | <b>13</b> |
| 2.1. Tổng quan về ReactJs .....                                         | 13        |
| 2.2. Tổng quan về ExpressJS.....                                        | 14        |
| 2.3. Tổng quan về FastAPI.....                                          | 16        |
| 2.4. Tổng quan về MongoDB .....                                         | 17        |
| 2.5. Tổng quan về Nginx .....                                           | 20        |
| 2.6. Tổng quan về Docker .....                                          | 21        |
| 2.7. Dịch vụ EC2 trên điện toán đám mây AWS .....                       | 22        |
| <b>CHƯƠNG 3. PHÂN TÍCH THIẾT KẾ HỆ THỐNG .....</b>                      | <b>23</b> |
| 3.1. Các tác nhân chính của hệ thống .....                              | 23        |
| 3.2. Biểu đồ usecase .....                                              | 23        |
| 3.2.1. Biểu đồ usecase tổng quan.....                                   | 23        |
| 3.2.2. Biểu đồ usecase phân rã.....                                     | 25        |
| 3.3. Đặc tả biểu đồ usecase.....                                        | 27        |
| 3.3.1. Tra từ.....                                                      | 27        |
| 3.3.2. Dịch câu .....                                                   | 28        |

|                                                          |    |
|----------------------------------------------------------|----|
| 3.3.3. Xem lịch sử tra từ .....                          | 28 |
| 3.3.4. Đăng nhập .....                                   | 29 |
| 3.3.5. Đăng ký .....                                     | 29 |
| 3.3.6. Quên mật khẩu .....                               | 30 |
| 3.3.7. Tạo bộ thẻ ghi nhớ .....                          | 30 |
| 3.3.8. Chỉnh sửa bộ thẻ ghi nhớ .....                    | 31 |
| 3.3.9. Xóa bộ thẻ ghi nhớ .....                          | 31 |
| 3.3.10. Tạo thẻ ghi nhớ .....                            | 32 |
| 3.3.11. Sửa thẻ ghi nhớ .....                            | 32 |
| 3.3.12. Xóa thẻ ghi nhớ .....                            | 32 |
| 3.3.13. Học thẻ ghi nhớ .....                            | 33 |
| 3.3.14. Thống kê học thẻ .....                           | 33 |
| 3.3.15. Chỉnh sửa thông tin cá nhân .....                | 34 |
| 3.3.16. Chỉnh sửa thông tin hiển thị .....               | 34 |
| 3.3.17. Thay đổi mật khẩu .....                          | 35 |
| 3.3.18. Thêm từ vựng .....                               | 35 |
| 3.3.19. Cập nhật từ vựng .....                           | 36 |
| 3.3.20. Xóa từ .....                                     | 37 |
| 3.3.21. Thêm hán tự .....                                | 37 |
| 3.3.22. Cập nhật hán tự .....                            | 38 |
| 3.3.23. Xem danh sách từ .....                           | 38 |
| 3.3.24. Xem danh sách người dùng .....                   | 39 |
| 3.3.25. Cập nhật thông tin tài khoản Quản trị viên ..... | 39 |
| 3.3.26. Xem thống kê .....                               | 40 |
| 3.4. Biểu đồ tuần tự .....                               | 40 |
| 3.4.1. Chức năng đăng nhập .....                         | 40 |
| 3.4.2. Chức năng đăng ký .....                           | 42 |
| 3.4.3. Chức năng tra cứu từ vựng .....                   | 42 |
| 3.4.4. Chức năng OCR hình ảnh .....                      | 43 |
| 3.4.5. Chức năng nhận diện chữ kanji .....               | 43 |
| 3.4.6. Chức năng dịch câu .....                          | 44 |
| 3.4.7. Chức năng detect giọng nói .....                  | 44 |
| 3.4.8. Chức năng học thẻ ghi nhớ .....                   | 45 |
| 3.4.9. Chức năng cập nhật thông tin người dùng .....     | 45 |
| 3.4.10. Chức năng thay đổi mật khẩu .....                | 45 |
| 3.4.11. Chức năng thống kê .....                         | 46 |
| 3.5. Biểu đồ hoạt động .....                             | 46 |
| 3.5.1. Chức năng đăng nhập .....                         | 46 |
| 3.5.2. Chức năng đăng ký .....                           | 47 |
| 3.5.3. Chức năng tra cứu từ vựng .....                   | 48 |

|                                                         |           |
|---------------------------------------------------------|-----------|
| 3.5.4. Chức năng dịch câu .....                         | 49        |
| 3.5.5. Chức năng học thẻ ghi nhớ .....                  | 50        |
| 3.5.6. Chức năng cập nhật thông tin người dùng .....    | 51        |
| 3.5.7. Chức năng thay đổi mật khẩu .....                | 52        |
| 3.6. Cơ sở dữ liệu .....                                | 53        |
| 3.6.1. Phân tích bảng cơ sở dữ liệu .....               | 53        |
| 3.6.1.1. User .....                                     | 53        |
| 3.6.1.2. Words .....                                    | 54        |
| 3.6.1.3. Kanjis .....                                   | 54        |
| 3.6.1.4. Compositions .....                             | 54        |
| 3.6.1.5. Decks .....                                    | 55        |
| 3.6.1.6. Cards .....                                    | 55        |
| 3.6.1.7. ReviewLogs .....                               | 56        |
| <b>CHƯƠNG 4. TRIỂN KHAI MÔ HÌNH VÀ THUẬT TOÁN .....</b> | <b>57</b> |
| 4.1. Phát biểu bài toán .....                           | 57        |
| 4.1.1. Dịch thuật .....                                 | 57        |
| 4.1.2. Thẻ ghi nhớ .....                                | 57        |
| 4.2. Triển khai mô hình và thuật toán .....             | 57        |
| 4.2.1. Triển khai mô hình Transformer .....             | 57        |
| 4.2.1.1. Cài đặt model .....                            | 57        |
| 4.2.1.2. Dữ liệu huấn luyện .....                       | 58        |
| 4.2.1.3. Tokenizer .....                                | 59        |
| 4.2.1.4. Huấn luyện model .....                         | 60        |
| 4.2.1.5. Kết quả và đánh giá mô hình .....              | 60        |
| 4.2.2. Triển khai thuật toán FSRS .....                 | 66        |
| <b>CHƯƠNG 5. TRIỂN KHAI THỰC TẾ .....</b>               | <b>69</b> |
| 5.1. Cấu trúc hệ thống .....                            | 69        |
| 5.2. Môi trường phát triển .....                        | 69        |
| 5.3. Môi trường triển khai .....                        | 69        |
| 5.4. Kết quả thực tế .....                              | 70        |
| 5.4.1. Giao diện người dùng .....                       | 70        |
| 5.4.2. Giao diện quản trị viên .....                    | 79        |
| <b>KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN .....</b>               | <b>84</b> |
| 1.1. Kết quả đạt được .....                             | 84        |
| 1.2. Những khó khăn và cách khắc phục .....             | 84        |
| 1.3. Những vấn đề còn tồn tại .....                     | 85        |
| 1.4. Hướng phát triển .....                             | 86        |
| <b>TÀI LIỆU THAM KHẢO .....</b>                         | <b>88</b> |

## DANH MỤC HÌNH ẢNH

|                                                                       |    |
|-----------------------------------------------------------------------|----|
| Hình 1 Kiến trúc Transformer .....                                    | 8  |
| Hình 2 Đường cong quên và sự phẳng dần qua từng lần gặp lại. ....     | 10 |
| Hình 3 Logo của ReacJS .....                                          | 13 |
| Hình 4 Logo của ExpressJS .....                                       | 14 |
| Hình 5 FastAPI.....                                                   | 16 |
| Hình 6 Logo MongoDB .....                                             | 18 |
| Hình 7 Nginx .....                                                    | 20 |
| Hình 8 Docker .....                                                   | 21 |
| Hình 9 Elastic Compute Cloud – Amazon EC2 .....                       | 22 |
| Hình 10 Các tác nhân trong hệ thống.....                              | 23 |
| Hình 11 Biểu đồ usecase tổng quan .....                               | 24 |
| Hình 12 Biểu đồ usecase phân rã theo “Khách” .....                    | 25 |
| Hình 13 Biểu đồ usecase phân rã theo “Người dùng” .....               | 26 |
| Hình 14 Biểu đồ usecase phân rã theo “Quản trị viên”.....             | 27 |
| Hình 15 Biểu đồ tuần tự chức năng đăng nhập.....                      | 41 |
| Hình 16 Biểu đồ tuần tự chức năng đăng ký.....                        | 42 |
| Hình 17 Biểu đồ tuần tự chức năng tra cứu từ vựng.....                | 42 |
| Hình 18 Biểu đồ tuần tự chức năng OCR hình ảnh .....                  | 43 |
| Hình 19 Biểu đồ tuần tự chức năng nhận diện chữ kanji.....            | 43 |
| Hình 20 Biểu đồ tuần tự chức năng dịch câu .....                      | 44 |
| Hình 21 Biểu đồ tuần tự chức năng detect giọng nói.....               | 44 |
| Hình 22 Biểu đồ tuần tự chức năng học thẻ ghi nhớ.....                | 45 |
| Hình 23 Biểu đồ tuần tự chức năng cập nhật thông tin người dùng ..... | 45 |
| Hình 24 Biểu đồ tuần tự chức năng thay đổi mật khẩu.....              | 46 |
| Hình 25 Biểu đồ tuần tự chức năng thống kê.....                       | 46 |
| Hình 26 Sơ đồ hoạt động chức năng đăng nhập.....                      | 46 |
| Hình 27 Sơ đồ hoạt động chức năng đăng ký .....                       | 47 |
| Hình 28 Sơ đồ hoạt động chức năng tra cứu từ vựng .....               | 48 |
| Hình 29 Sơ đồ hoạt động chức năng Dịch câu.....                       | 49 |
| Hình 30 Sơ đồ hoạt động chức năng Học thẻ ghi nhớ .....               | 50 |
| Hình 31 Sơ đồ hoạt động chức năng Cập nhật thông tin người dùng.....  | 51 |
| Hình 32 Sơ đồ hoạt động chức năng Thay đổi mật khẩu .....             | 52 |
| Hình 33 Cơ sở dữ liệu .....                                           | 53 |
| Hình 34 Tổng quan dataset.....                                        | 59 |
| Hình 35 Xử lý tokenizer.....                                          | 60 |
| Hình 36 Hoạt động sử dụng của thuật toán FSRS .....                   | 66 |
| Hình 37 Tổng quan về hệ thống .....                                   | 69 |

|                                                        |    |
|--------------------------------------------------------|----|
| Hình 38 Giao diện tra cứu từ vựng.....                 | 71 |
| Hình 39 Giao diện dịch .....                           | 72 |
| Hình 40 Giao diện đăng nhập.....                       | 73 |
| Hình 41 Giao diện đăng ký .....                        | 74 |
| Hình 42 Giao diện bộ thẻ .....                         | 75 |
| Hình 43 Giao diện modal thêm thẻ vào bộ thẻ.....       | 76 |
| Hình 44 Giao diện học thẻ.....                         | 77 |
| Hình 45 Giao diện thống kê thẻ học.....                | 77 |
| Hình 46 Giao diện cập nhật thông tin người dùng.....   | 78 |
| Hình 47 Giao diện cập nhật mật khẩu.....               | 78 |
| Hình 48 Giao diện thống kê hệ thống .....              | 79 |
| Hình 49 Giao diện danh sách từ vựng.....               | 79 |
| Hình 50 Giao diện modal thêm và cập nhật từ vựng ..... | 80 |
| Hình 51 Giao diện danh sách hán tự .....               | 81 |
| Hình 52 Giao diện modal thêm và cập nhật hán tự.....   | 82 |
| Hình 53 Giao diện danh sách người dùng.....            | 83 |

## DANH MỤC BẢNG BIỂU

|                                                                         |    |
|-------------------------------------------------------------------------|----|
| Bảng 1 Các chức năng chính của hệ thống .....                           | 23 |
| Bảng 2 Đặc tả usecase tra từ .....                                      | 27 |
| Bảng 3 Đặc tả usecase dịch câu .....                                    | 28 |
| Bảng 4 Đặc tả usecase xem lịch sử tra từ.....                           | 28 |
| Bảng 5 Đặc tả usecase đăng nhập .....                                   | 29 |
| Bảng 6. Đặc tả usecase đăng ký .....                                    | 30 |
| Bảng 7. Đặc tả usecase quên mật khẩu .....                              | 30 |
| Bảng 8 Đặc tả usecase tạo bộ thẻ ghi nhớ.....                           | 31 |
| Bảng 9 Đặc tả usecase chỉnh sửa bộ thẻ ghi nhớ .....                    | 31 |
| Bảng 10 Đặc tả usecase xóa bộ thẻ ghi nhớ .....                         | 31 |
| Bảng 11 Đặc tả usecase tạo thẻ ghi nhớ .....                            | 32 |
| Bảng 12 Đặc tả usecase sửa thẻ ghi nhớ .....                            | 32 |
| Bảng 13 Đặc tả usecase xóa thẻ ghi nhớ .....                            | 33 |
| Bảng 14 Đặc tả usecase học thẻ ghi nhớ .....                            | 33 |
| Bảng 15 Đặc tả usecase thống kê học thẻ .....                           | 34 |
| Bảng 16 Đặc tả usecase chỉnh sửa thông tin cá nhân.....                 | 34 |
| Bảng 17 Đặc tả usecase chỉnh sửa thông tin hiển thị .....               | 35 |
| Bảng 18 Đặc tả usecase thay đổi mật khẩu .....                          | 35 |
| Bảng 19 Đặc tả usecase thêm từ vựng .....                               | 36 |
| Bảng 20 Đặc tả usecase cập nhật từ vựng .....                           | 36 |
| Bảng 21 Đặc tả usecase xóa từ.....                                      | 37 |
| Bảng 22 Đặc tả usecase thêm hán tự.....                                 | 38 |
| Bảng 23 Đặc tả usecase cập nhật hán tự .....                            | 38 |
| Bảng 24 Đặc tả usecase xem danh sách từ.....                            | 39 |
| Bảng 25 Đặc tả usecase xem danh sách từ.....                            | 39 |
| Bảng 26 Đặc tả usecase cập nhật thông tin tài khoản Quản trị viên ..... | 40 |
| Bảng 27 Đặc tả usecase xem thống kê .....                               | 40 |
| Bảng 28 Phân tích bảng User .....                                       | 54 |
| Bảng 29 Phân tích bảng Words .....                                      | 54 |
| Bảng 30 Phân tích bảng Kanjis .....                                     | 54 |
| Bảng 31 Phân tích bảng Compositions .....                               | 54 |
| Bảng 32 Phân tích bảng Decks.....                                       | 55 |
| Bảng 33 Phân tích bảng Cards .....                                      | 55 |
| Bảng 34 Phân tích bảng ReviewLogs .....                                 | 56 |

## MỞ ĐẦU

### 1. Giới thiệu

Ứng dụng từ điển tiếng Nhật thông minh tích hợp AI và phương pháp ghi nhớ khoa học là một dự án được thực hiện với mục đích tạo ra công cụ hỗ trợ học và tra cứu ngôn ngữ hiệu quả hơn, đáp ứng nhu cầu ngày càng cao của sinh viên, nhân viên văn phòng và những người đam mê tiếng Nhật. Trong bối cảnh nền kinh tế – văn hóa Nhật Bản ngày càng sâu rộng, việc nắm vững từ vựng, cấu trúc câu và khả năng dịch chính xác trở thành yếu tố quan trọng để giao tiếp và tiếp cận tài liệu gốc. Tuy nhiên, các ứng dụng từ điển truyền thống vẫn còn hạn chế ở việc chỉ tra nghĩa từ đơn lẻ, chưa chú trọng đến ngữ cảnh, chưa kết hợp phương pháp học tập khoa học để giúp người dùng ghi nhớ từ vựng lâu dài.

Dự án này được xây dựng nhằm giải quyết các thách thức trên thông qua việc tích hợp mô hình dịch AI giúp dịch nguyên câu theo ngữ nghĩa tự nhiên, thay vì dịch từng từ rời rạc, giúp người dùng hiểu đúng và cảm nhận được sắc thái ngôn ngữ. Ngoài ra, ứng dụng còn hỗ trợ tính năng OCR (Optical Character Recognition), cho phép người dùng chụp ảnh văn bản tiếng Nhật – từ sách, báo, và ngay lập tức nhận được bản dịch. Nhờ vậy, việc tra cứu và tiếp cận nội dung gốc trở nên linh hoạt và thuận tiện hơn.

Để đảm bảo khả năng ghi nhớ từ vựng và ôn tập hiệu quả, ứng dụng áp dụng hệ thống “Forgotten Spaced Repetition System” (FSRS) – một phương pháp ghi nhớ ngắt quãng đã được chứng minh giúp tối ưu hóa quá trình lưu trữ thông tin vào bộ nhớ dài hạn. Hệ thống này sẽ tự động lên lịch nhắc nhở ôn tập dựa trên mức độ ghi nhớ thực tế của người dùng, giúp tránh tình trạng “học trước quên sau” và nâng cao hiệu quả học tập. Đồng thời, người dùng có thể tự tạo bộ thẻ từ vựng cá nhân hóa, đánh dấu từ khó, theo dõi tiến độ học tập, nhận gợi ý từ vựng cần ôn tập tiếp theo dựa trên kết quả đánh giá mức độ nhớ từ.

Em tin rằng, với sự kết hợp giữa công nghệ AI, OCR và phương pháp FSRS, ứng dụng từ điển tiếng Nhật thông minh này sẽ mang đến cho người học một môi trường học tập toàn diện, linh hoạt và hiệu quả. Thông qua ứng dụng, em hy vọng sẽ góp phần tạo ra công cụ hỗ trợ đắc lực, giúp cộng đồng học tiếng Nhật nâng cao trình độ nhanh chóng và bền vững.

### 2. Mục đích và ý nghĩa đề tài

#### 2.1. Mục đích

Mục đích chính của ứng dụng từ điển tiếng Nhật thông minh là xây dựng một công cụ hỗ trợ tra cứu và học từ vựng hiệu quả, đáp ứng nhu cầu ngày càng cao của sinh viên, người đi làm và những ai đam mê tiếng Nhật. Ứng dụng không chỉ đơn thuần cung cấp nghĩa của từ mà còn giúp người dùng hiểu đúng ngữ cảnh nhờ mô hình dịch AI chuyển nguyên câu theo ngữ nghĩa tự nhiên, giúp việc tiếp cận văn bản gốc trở nên chính xác và dễ dàng hơn.

Bên cạnh đó, ứng dụng tích hợp công nghệ OCR để người dùng có thể chụp ảnh văn bản tiếng Nhật và ngay lập tức nhận được bản dịch cùng giải thích chi tiết. Tính năng này giúp việc tra cứu tiện lợi khi không có sẵn tài liệu số, tiết kiệm thời gian và hỗ trợ học tập linh hoạt mọi lúc, mọi nơi.

Để đảm bảo quá trình ghi nhớ từ vựng lâu dài, ứng dụng áp dụng hệ thống FSRS (Forgotten Spaced Repetition System), tự động lên lịch nhắc nhở ôn tập dựa trên khả năng ghi nhớ của từng cá nhân. Người dùng có thể tạo bộ thẻ từ vựng cá nhân hóa, đánh dấu từ khó và theo dõi tiến độ học; từ đó, ứng dụng sẽ đề xuất danh sách từ cần ôn tập tiếp theo, giúp tối ưu hóa thời gian ôn luyện và tránh tình trạng “học trước quên sau”.

## 2.2. Ý nghĩa

Việc tích hợp AI dịch ngữ nghĩa vào ứng dụng giúp chuyển đổi câu tiếng Nhật sang tiếng Việt (và ngược lại) một cách tự nhiên, chính xác hơn, hạn chế tình trạng “dịch cứng” từng từ. Nhờ vậy, người dùng dễ dàng nắm bắt được nội dung và sắc thái của câu gốc mà không mất nhiều thời gian tra cứu từng thành phần. Bên cạnh đó, thuật toán FSRS (Forgotten Spaced Repetition System) đảm bảo mỗi từ vựng được lặp lại theo khoảng thời gian tối ưu, vừa tránh học trước quên sau, vừa tăng khả năng ghi nhớ dài hạn. Kết hợp hai công nghệ này, ứng dụng không chỉ là công cụ tra cứu đơn thuần mà còn hỗ trợ người học ôn tập hiệu quả, tiết kiệm công sức và tạo động lực duy trì việc học hàng ngày.

## 3. Các bước triển khai

Giai đoạn 1: Xây dựng cơ sở dữ liệu và các biểu đồ liên quan đến đề tài, bao gồm việc thiết kế cấu trúc cơ sở dữ liệu cho các thông tin về từ vựng, log review thẻ ghi nhớ, kanji và các thông tin liên quan. Thiết kế giao diện người dùng.

Giai đoạn 2: Xây dựng giao diện người dùng và lập trình các chức năng tra cứu từ vựng, chức năng ocr tài liệu, chức năng detect giọng nói, chức năng học thẻ theo phương pháp ghi nhớ ngắt quãng và các chức năng quản lý người dùng khác.

Giai đoạn 3: Tiến hành tìm hiểu về model training dịch thuật ngôn ngữ Nhật-Việt và Việt-Nhật. Chuẩn bị dữ liệu training. Thực hiện training model và điều chỉnh

thông số để đạt được model training tốt nhất. Áp dụng model dịch thuật xây dựng chức năng dịch cho ứng dụng.

Giai đoạn 4: Viết báo cáo và chạy song song.

#### **4. Công nghệ và kỹ thuật sử dụng**

Các công nghệ và kỹ thuật được sử dụng

- Front-end: ReactJs.
- Back-end: ExpressJs.
- Database: MongoDB.
- Quản lý mã nguồn: Git, Github.
- Server AI: FastAPI.
- Công cụ: Visual Studio Code, Postman.

#### **5. Những kết quả đạt được dự kiến**

##### **5.1. Về mặt lý thuyết**

- Hiểu được quy trình phát triển và xây dựng phần mềm.
- Vận dụng các kiến thức đã được học ở trường và trong quá trình thực tập tốt nghiệp để phân tích và thiết kế một hệ thống.
- Hiểu và thực hiện được các bước về thiết kế cơ sở dữ liệu, thiết kế API và giao diện người dùng.

##### **5.2. Về mặt ứng dụng**

- Xây dựng ứng dụng từ điển tiếng Nhật thông minh: Ứng dụng được phát triển nhằm cung cấp nền tảng tra cứu và dịch toàn diện, sử dụng mô hình AI để dịch nguyên câu theo ngữ nghĩa tự nhiên, giúp người dùng hiểu đúng hoàn cảnh sử dụng từ.
- Tích hợp tính năng OCR (Optical Character Recognition): Người dùng có thể upload ảnh văn bản tiếng Nhật và ngay lập tức nhận được bản dịch, phiên âm và giải thích chi tiết, giúp tra cứu tiện lợi khi không có sẵn tài liệu số.
- Áp dụng hệ thống FSRS để ôn tập từ vựng: Ứng dụng tự động lên lịch nhắc nhở ôn tập dựa trên mức độ ghi nhớ thực tế, tránh tình trạng “học trước quên sau”. Người học có thể tạo bộ thẻ từ cá nhân hóa, đánh dấu từ khó, theo dõi tiến độ và nhận gợi ý từ vựng cần ôn tập tiếp theo.
- Dễ dàng tra cứu thông tin và tiện ích cho người học: Giao diện thân thiện, các chức năng tra nghĩa, dịch câu, tạo flashcards và theo dõi tiến trình học tập đều được tích hợp tại một nơi. Người dùng chỉ cần vài thao tác đơn giản để truy cập và sử dụng mọi lúc, mọi nơi.

- Tăng tính minh bạch và an toàn cho người dùng: Ứng dụng hoạt động trên nền tảng ổn định, đảm bảo bảo mật dữ liệu cá nhân và kết quả học tập. Mọi thông tin tra cứu, ghi nhớ và tiến độ ôn tập đều được lưu trữ an toàn, tạo môi trường học tập đáng tin cậy.

## 6. Nội dung đồ án

Đồ án bao gồm nội dung sau:

- **Mở đầu** - Giới thiệu tổng quan, mục đích và ý nghĩa của đề tài, công nghệ, kỹ thuật, công cụ được sử dụng và những kết quả dự kiến đạt được.
- **Chương 1: Cơ sở lý thuyết** – Trình bày những cơ sở lý thuyết được áp dụng trong đề tài.
- **Chương 2: Công cụ phát triển** – Trình bày cơ sở lý thuyết các công cụ phát triển được sử dụng trong hệ thống.
- **Chương 3: Phân tích thiết kế hệ thống** – Phân tích các yêu cầu về chức năng và phi chức năng của hệ thống và triển khai thiết kế hệ thống thông qua các biểu đồ thiết kế hệ thống.
- **Chương 4: Triển khai mô hình và thuật toán** - Trình bày quá trình triển khai training mô hình và sử dụng thuật toán trong ứng dụng.
- **Chương 5: Triển khai thực tế** - Trình bày môi trường triển khai và kết quả thực tế đạt được.
- **Kết luận và hướng phát triển** – Đánh giá kết quả đạt được, chưa đạt được và đưa ra những định hướng phát triển thêm trong tương lai.
- **Tài liệu tham khảo** – Liệt kê các tài liệu tham khảo sử dụng trong đề tài.

## CHƯƠNG 1. CƠ SỞ LÝ THUYẾT

### 1.1. Tổng quan về mô hình Transformer

Transformer là một kiến trúc mạng nơ-ron sâu được giới thiệu lần đầu bởi Vaswani và cộng sự trong bài báo “Attention is All You Need” (2017), mở ra một bước ngoặt lớn trong lĩnh vực xử lý ngôn ngữ tự nhiên (NLP). Khác với các mô hình tuần tự trước đó như mạng nơ-ron hồi quy (RNNs), bao gồm cả Long Short-Term Memory (LSTM) vốn phụ thuộc vào việc xử lý dữ liệu theo chuỗi từ đầu đến cuối, Transformer hoàn toàn loại bỏ tính tuần tự trong kiến trúc, thay vào đó sử dụng cơ chế **self-attention** để xử lý dữ liệu một cách toàn cục, hiệu quả và song song hóa cao.

#### 1.1.1. Vấn đề của mô hình tuần tự (RNN, LSTM)

Trong các mô hình truyền thống như RNN hoặc LSTM, dữ liệu đầu vào được xử lý theo từng bước một, từ trái sang phải. Cụ thể, mỗi từ trong câu chỉ được xử lý sau khi từ trước nó đã được xử lý xong. Điều này khiến cho quá trình huấn luyện không thể song song hóa, vì mỗi bước phụ thuộc vào bước trước đó. Mặt khác, các mô hình này còn gặp khó khăn khi xử lý các quan hệ dài hạn trong câu, ví dụ như mối quan hệ giữa chủ ngữ và động từ cách xa nhau.

Ví dụ, trong câu:

"The book that the professor assigned was very difficult."

Động từ "was" phải liên kết đúng với chủ ngữ "book", không phải "professor". Với RNN, thông tin này dễ bị mờ nhạt do độ dài câu tăng.

#### 1.1.2. Cách Transformer xử lý toàn cục bằng Self-Attention

Transformer xử lý dữ liệu theo chuỗi nhưng không theo cách tuần tự, mà mỗi từ trong chuỗi sẽ được xử lý cùng lúc với tất cả các từ khác. Điều này được thực hiện thông qua cơ chế self-attention, nơi mà mỗi từ có thể nhìn toàn bộ câu để xác định xem mình nên tập trung vào từ nào nhiều hơn.

Ví dụ, với câu: "The animal didn't cross the street because it was too tired", để hiểu "it" là gì, mô hình cần tập trung (attend) vào "animal". Cơ chế self-attention giúp mô hình làm được điều đó bằng cách tính toán mức độ liên kết giữa "it" và tất cả các từ khác trong câu.

Việc xử lý toàn cục này còn có thể thực hiện song song cho toàn bộ câu nhờ việc tất cả các từ đều có thể tính toán cùng lúc. Khác với RNN vốn cần đợi từng bước một, Transformer cho phép tính toán nhanh hơn nhiều, đặc biệt khi áp dụng trên GPU/TPU.

#### 1.1.3. Kiến trúc tổng thể

Cốt lõi của transformer là attention mechanism (cơ chế tập trung), giúp mô hình

tập trung vào các phần quan trọng của văn bản để đưa ra dự đoán chính xác hơn. **Transformer** được cấu trúc thành hai phần chính là encoder và decoder.

- **Encoder:** Encoder xử lý dữ liệu đầu vào (gọi là "Source") và nén dữ liệu vào vùng nhớ hoặc context mà Decoder có thể sử dụng sau đó.
- **Decoder:** Decoder nhận đầu vào từ đầu ra của Encoder (gọi là "Encoded input") kết hợp với một chuỗi đầu vào khác (gọi là "Target") để tạo ra chuỗi đầu ra cuối cùng.

Mỗi encoder và decoder đều bao gồm nhiều lớp, mỗi lớp chứa các thành phần self-attention và feed-forward neural networks. Về cơ bản, quy trình hoạt động của Transformer gồm các bước sau:

### Bước 1: Preprocessing (Tiền xử lý)

1. Tokenization: Dữ liệu đầu vào được biến đổi thành token.
2. Embedding: Biến token thành vector. Mỗi token được chuyển thành một vector.
3. Positional Encoding: Vì Transformer không xử lý tuần tự nên cần một cách để hiểu vị trí của từng từ trong câu. Điều này được thực hiện thông qua positional encodings, được cộng trực tiếp vào input embeddings. Các positional encodings có thể dùng các hàm sin và cos với bước sóng khác nhau để mỗi vị trí có một encoding duy nhất.

### Bước 2: Encoder

Encoder bao gồm nhiều layer (lớp) xếp chồng lên nhau (stack). Mỗi layer có 2 sublayer là **Multi-Head Attention** và **Positionwise Feed-forward Neural Network (FNN)**. Output của 2 sublayer sẽ đi qua một lớp gọi là **Add & Norm**.

- **Multi-Head Attention:** Dữ liệu đầu vào được truyền vào nhiều head để tạo một tập các vector đầu ra. Multi-head attention được thiết kế để cho phép mô hình xử lý thông tin đồng thời ở các không gian (subspaces) khác nhau. Thay vì có một module attention duy nhất, mô hình có nhiều "heads", mỗi head có thể tập trung vào các phần khác nhau của đầu vào.
- **Positionwise Feed-forward Neural Network (FNN):** Các vector đầu ra của attention head được truyền qua positionwise feed-forward neural network (mạng thần kinh chuyển tiếp theo vị trí). Mỗi head trong khối có một mạng FNN riêng, cho phép mô hình học được các biểu diễn dựa trên nhiều không gian (subspaces) khác nhau của dữ liệu. Mục tiêu là tạo một lớp biểu diễn phi tuyến tính (non-linear representation) của dữ liệu đầu vào, giúp mô hình có khả năng học được các mối quan hệ phức tạp và phi tuyến tính của dữ liệu.
- **Add & Norm:** Bao gồm hai bước Add và Norm. Bước Add thêm residual connection (kết nối dư) nhằm giảm vấn đề vanishing gradient (mất độ dốc) trong các

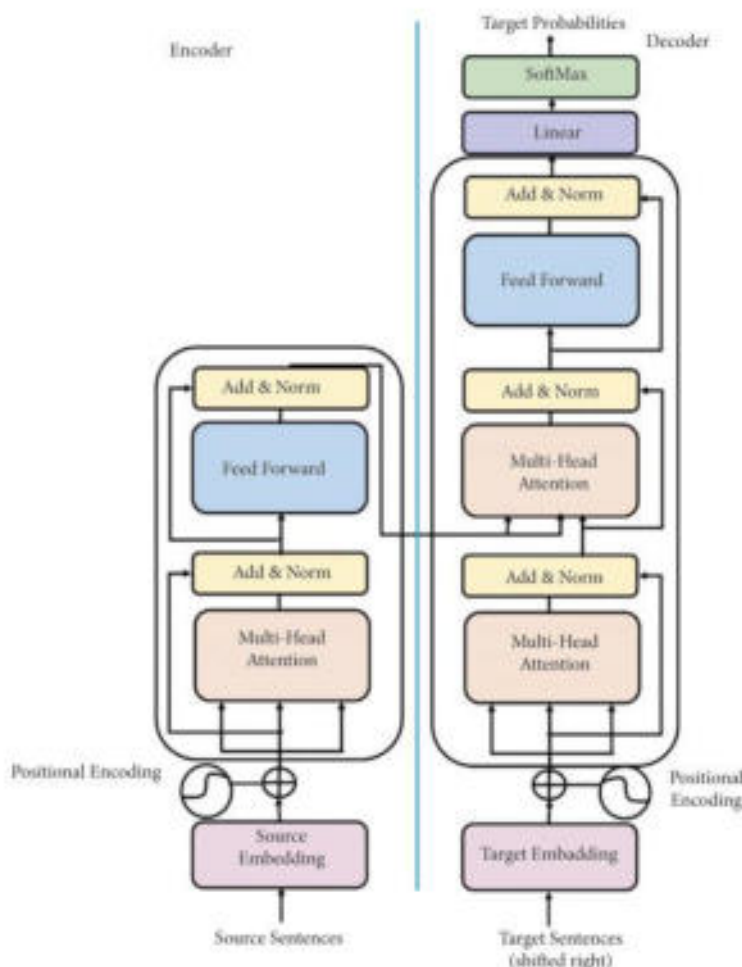
deep network (mạng sâu). Ngay sau đó, bước Norm thực hiện layer normalization (chuẩn hoá lớp) giúp ổn định quá trình huấn luyện và giảm số lượng giai đoạn cần thiết để huấn luyện.

### Bước 3: Decoder

Tương tự Encoder, Decoder cũng có 2 sublayer và có thêm 1 sublayer ở nằm giữa 2 sublayer tên là Encode-decode Attention.

- **Masked Multi-Head Attention:** Là một biến thể của Multi-Head Attention, hoạt động tương tự Multi-Head Attention nhưng có "mask" để đảm bảo rằng các vị trí trong output chỉ có thể tập trung vào các từ trước đó mà không "nhìn thấy" các từ tiếp theo (điều này quan trọng cho quá trình dự đoán từ tiếp theo).
- **Multi-Head Attention:** Hoạt động giống với bước Encoder.
- **Encode-decode Attention:** Trong Encode-decode Attention, truy vấn tới từ đầu ra của lớp con Attention, còn key và value tới từ đầu ra của Encoder.
- **Positionwise Feed-forward Neural Network (FNN):** Hoạt động giống với bước Encoder.
- **Add & Norm:** Hoạt động giống với bước Encoder.

Đầu ra của Decoder thường là một chuỗi các vector. Mỗi vector tương ứng với một từ trong chuỗi đầu ra. Các vector này sau đó được chuyển qua một lớp tuyến tính (linear layer) và một hàm softmax để tạo ra một phân phối xác suất trên tất cả các từ trong từ điển. Sau cùng, từ có xác suất cao nhất được chọn làm từ dự đoán tiếp theo trong chuỗi. Quy trình này lặp lại cho mỗi từ trong chuỗi đầu ra, cho tới khi mô hình dự đoán ra ký tự kết thúc câu hoặc đạt độ dài tối đa cho phép.



Hình 1 Kiến trúc Transformer

#### 1.1.4. Cơ chế Attention và Multi-head Attention

##### Self-Attention hoạt động như thế nào?

Giả sử chuỗi đầu vào có các từ: "The cat sat on the mat"

Đối với mỗi từ, mô hình tạo ra 3 vector:

- Query (Q): đại diện cho câu hỏi: "Tôi nên tập trung vào từ nào?"
- Key (K): đại diện cho nội dung của từng từ còn lại.
- Value (V): thông tin của từ đó để truyền đi nếu được chú ý.

Mô hình tính điểm attention bằng cách so sánh Q của từ hiện tại với K của tất cả các từ khác. Từ đó mô hình biết nên "chú ý" vào từ nào nhiều hay ít. Cuối cùng, các điểm attention này được dùng để tạo tổ hợp tuyến tính của các vector V tương ứng.

Ví dụ: Từ "sat" có thể chú ý nhiều hơn đến "cat" để hiểu ai là người đang ngồi.

##### Multi-head Attention

Thay vì chỉ có một cặp Q-K-V duy nhất, Multi-head Attention tách không gian

biểu diễn thành nhiều phần và thực hiện self-attention song song trên mỗi phần này.

Nhờ đó, mô hình có thể học được nhiều loại quan hệ khác nhau trong cùng một lớp attention — chẳng hạn một đầu chú ý ngữ nghĩa, một đầu chú ý cấu trúc câu, v.v.

#### 1.1.5. Triển khai và ứng dụng

Transformer được huấn luyện đầu tiên để thực hiện nhiệm vụ dịch máy (machine translation). Ví dụ, từ chuỗi đầu vào tiếng Bồ Đào Nha, mô hình học cách sinh ra chuỗi đầu ra tiếng Anh. Sau đó, kiến trúc này được mở rộng thành các mô hình tiền huấn luyện (pre-trained) như BERT, GPT, T5 và thậm chí ứng dụng trong thị giác máy tính như Vision Transformer (ViT).

#### 1.1.6. Ưu điểm

- Song song hóa hiệu quả: Không cần xử lý tuần tự như RNN → tăng tốc độ huấn luyện rõ rệt.
- Nắm bắt ngữ cảnh dài hạn tốt hơn: attention có thể liên kết từ bất kỳ vị trí nào với nhau.
- Khả năng mở rộng cao: Dễ dàng tăng số lớp hoặc số đầu attention.
- Đa năng: Tùy biến thành encoder-only (BERT), decoder-only (GPT), encoder-decoder (T5).

#### 1.1.7. Hạn chế

Dù rất mạnh mẽ, Transformer vẫn có những hạn chế:

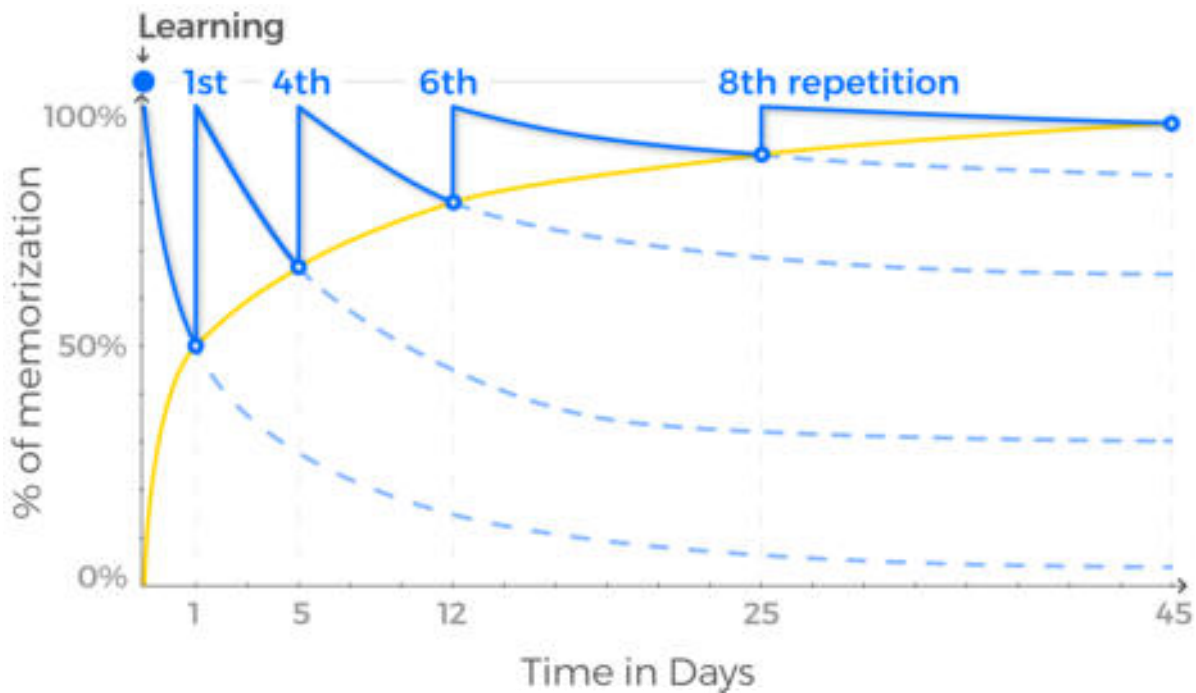
- **Chi phí tính toán cao:** Attention có độ phức tạp  $O(n^2)$  theo độ dài chuỗi, gây tốn tài nguyên với dữ liệu dài.
- **Cần nhiều dữ liệu huấn luyện:** Để phát huy toàn bộ sức mạnh, mô hình thường cần lượng lớn dữ liệu và GPU mạnh.
- **Thiếu hiểu biết cấu trúc:** Không tận dụng thông tin ngữ pháp rõ ràng như các cây cú pháp (parse tree).

### 1.2. Tổng quan về thuật toán FSRS

**FSRS (Free Spaced Repetition Scheduler)** là một thuật toán lập lại cách quãng (spaced repetition) hiện đại, được phát triển để tối ưu hóa quá trình ghi nhớ lâu dài. Khác với các thuật toán truyền thống như SM-2, FSRS dựa trên mô hình toán học về trí nhớ, giúp dự đoán khả năng quên và điều chỉnh lịch ôn tập một cách linh hoạt, cá nhân hóa và hiệu quả hơn.

#### 1.2.1. Đường cong quên và sự phẳng dần qua mỗi lần lặp.

**Đường cong quên:**



Hình 2 Đường cong quên và sự phẳng dần qua từng lần gặp lại.

Đường màu xanh biểu diễn khả năng nhớ lại chỉ một lần học ban đầu (gần như ngay sau khi học, khả năng nhớ = 100%). Khả năng này giảm nhanh dần theo thời gian — hiện tượng gọi là “đường cong quên”

Các đường màu vàng thể hiện khả năng nhớ lại sau mỗi lần lặp lại:

- Mỗi lần ôn (điểm “đỉnh” trên đồ thị), khả năng nhớ lại được phục hồi lên gần 100%.
- Sau đó, đường cong trở nên phẳng hơn, nghĩa là tốc độ quên chậm đi sau mỗi lần ôn.

Ý nghĩa:

- Mỗi lần học lại, khả năng nhớ lại được củng cố kỹ hơn.
- Khoảng thời gian ôn lại (giữa các lần xuất hiện “đỉnh”) có thể kéo dài hơn, chi phí học thấp hơn, nhưng vẫn đảm bảo retention cao.

**Sự phẳng dần qua mỗi lần lặp lại (hiệu ứng của spaced repetition):**

- Nhiều đường cong sau mỗi lần ôn dần thu nhỏ độ dốc (đường “phẳng dần”).
- Sau mỗi lần ôn, bạn không chỉ nằm trên đỉnh mà còn được “neo” ở vị trí cao hơn so với lần trước, thể hiện qua đường cong dịch lên.

Kết quả thực tế:

- Sau một lần ôn, bạn có thể ôn lại với khoảng thời gian dài hơn so với lần trước.

- Sự dày đặc ban đầu (lặp lại sau 1–2 ngày) giúp củng cố, sau đó mỗi lần cách xa hơn (5 ngày, 10 ngày, 30 ngày...), duy trì retention lâu dài.

### 1.2.2. Nguyên lý hoạt động của FSRS và áp dụng vào đường cong quên.

Mô hình DSR gồm ba thành phần cơ bản:

1. **Difficulty (D)** – Độ khó của thẻ (flashcard). Thẻ càng khó thì cần nhiều lần ôn hơn.
2. **Stability (S)** – Độ bền của trí nhớ. Độ bền cao nghĩa là thông tin được lưu giữ lâu hơn.
3. **Retrievability (R)** – Xác suất nhớ lại thẻ tại thời điểm ôn.

**Thuật toán FSRS sử dụng mô hình DSR để:**

1. Tính “Stability” (S) hiện tại dựa trên lịch sử ôn và phản hồi của người dùng.
2. Dự đoán thời điểm ôn kế tiếp sao cho khả năng nhớ (Retrievability) đạt mục tiêu, ví dụ ~90%.
3. Cập nhật “Stability” tiếp để đường cong quên ngày càng phẳng hơn.

Ví dụ:

- Lần đầu: S nhỏ → đường quên dốc → ôn sớm (1–2 ngày).
- Sau vài lần ôn: S lớn hơn → đường quên phẳng → ôn cách ngày dài hơn (5–10 ngày...).

**Quy trình FSRS gồm các bước:**

1. Đánh giá khi ôn: Người dùng tự đánh giá mức độ nhớ lại của thẻ theo 4 mức:
  - Again (0): Quên hoàn toàn.
  - Hard (1): Khó nhớ.
  - Good (2): Nhớ vừa phải.
  - Easy (3): Nhớ rất dễ.
2. Cập nhật thông số: Dựa trên phản hồi, thuật toán điều chỉnh:
  - “Stability” theo công thức cập nhật.
  - “Difficulty” theo mức độ đánh giá.
3. Tính ngày ôn tiếp theo: Chọn khoảng thời gian  $t_{next}$  sao cho

$$R(t_{next}) = R^* \text{ (} R^*: \text{Khả năng nhớ lại mong muốn)}$$

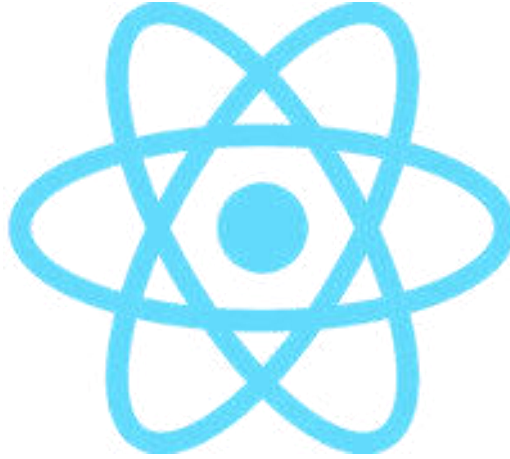
Suy ra khoảng thời gian:

$$t_{next} = -S_{new} \cdot \ln(R^*)$$

4. Áp dụng fuzzing (tùy chọn): Thêm nhiễu nhỏ vào để tránh tập trung ôn quá nhiều thẻ cùng lúc.

## CHƯƠNG 2. CÔNG CỤ PHÁT TRIỂN

### 2.1. Tổng quan về ReactJs



Hình 3 Logo của ReactJS

ReactJS là thư viện JavaScript được sử dụng để xây dựng các thành phần UI có thể tái sử dụng. React được sử dụng tại Facebook trong production, và instagram được viết hoàn toàn trên React.

Một trong những điểm hấp dẫn của React là thư viện này không chỉ hoạt động trên phía client, mà còn được render trên server và có thể kết nối với nhau. React so sánh sự thay đổi giữa các giá trị của lần render này với lần render trước và cập nhật ít thay đổi nhất trên DOM. React cũng có thể hiển thị trên máy chủ bằng Node và nó có thể cung cấp năng lượng cho các ứng dụng gốc bằng React Native.

ReactJS có nhiều tính năng hữu ích cho việc phát triển ứng dụng web, bao gồm:

- **Components:** ReactJS cho phép phát triển ứng dụng web theo mô hình component. Các component là các phần tử UI độc lập có thể được tái sử dụng trong nhiều phần khác nhau của ứng dụng.
- **Virtual DOM:** ReactJS sử dụng Virtual DOM để tối ưu hóa hiệu suất của ứng dụng. Virtual DOM là một bản sao của DOM được lưu trữ trong bộ nhớ và được cập nhật một cách nhanh chóng khi có thay đổi, giúp tăng tốc độ và hiệu suất của ứng dụng.
- **JSX:** JSX là một ngôn ngữ lập trình phân biệt được sử dụng trong ReactJS để mô tả các thành phần UI. JSX kết hợp HTML và JavaScript, giúp cho việc viết mã dễ hiểu và dễ bảo trì hơn.
- **State và Props:** ReactJS cho phép quản lý trạng thái của các thành phần UI thông qua State và Props. State là trạng thái của một thành phần được quản lý bởi nó

chính, trong khi Props là các giá trị được truyền vào từ bên ngoài để tùy chỉnh hoặc điều khiển hành vi của một thành phần.

- **Hỗ trợ tốt cho SEO:** ReactJS hỗ trợ tốt cho việc tối ưu hóa SEO. Với các thư viện như React Helmet, các nhà phát triển có thể quản lý các phần tử meta và title cho từng trang web, giúp tăng khả năng tìm kiếm và tăng cường trải nghiệm người dùng.
- **Hỗ trợ đa nền tảng:** ReactJS không chỉ được sử dụng để phát triển ứng dụng web, mà còn được sử dụng để phát triển ứng dụng di động với React Native. Sử dụng React Native, các nhà phát triển có thể xây dựng ứng dụng di động cho cả iOS và Android sử dụng cùng một mã nguồn.
- **Redux:** Redux là một thư viện quản lý trạng thái cho các ứng dụng ReactJS. Nó giúp quản lý trạng thái của ứng dụng một cách chính xác và dễ dàng, đồng thời giúp tăng tính linh hoạt và khả năng mở rộng của ứng dụng.

## 2.2. Tổng quan về ExpressJS



Hình 4 Logo của ExpressJS

Express.js là một framework web nhẹ, linh hoạt và mạnh mẽ dành cho nền tảng Node.js, được sử dụng rộng rãi để xây dựng các ứng dụng backend, đặc biệt là các RESTful API. Trong đề tài “Xây dựng website từ điển Việt Nhật có ứng dụng mô hình AI dịch thuật”, Express.js được chọn làm server chính để xử lý các yêu cầu từ phía client, quản lý luồng dữ liệu và kết nối với cơ sở dữ liệu. Dưới đây là những điểm nổi bật của Express.js:

- **Kiến trúc tối giản, tập trung vào middleware & routing:**
  - Express.js cung cấp hệ thống middleware linh hoạt, cho phép xử lý tuần tự nhiều chức năng (như xác thực, phân tích body, logging...) trước khi đến bước xử lý chính.
  - Cơ chế routing của Express giúp định nghĩa rõ ràng các endpoint (ví

dụ: GET /api/v1/translate,...) và dễ dàng tổ chức mã nguồn theo từng module.

- **Xây dựng RESTful API nhanh chóng và hiệu quả:**

- Với cú pháp đơn giản (app.get(), app.post(), app.put(), app.delete()), lập trình viên có thể thiết lập các route phục vụ nhu cầu tra cứu từ vựng, dịch câu, thêm thẻ học, quản lý người dùng... một cách gọn nhẹ.
- Express.js hỗ trợ tốt việc xử lý JSON, cho phép client gửi và nhận dữ liệu ở định dạng tiêu chuẩn, rất phù hợp khi tích hợp với frontend React/Vue hoặc các ứng dụng di động.

- **Khả năng mở rộng linh hoạt:**

- Bằng cách sử dụng middleware, ứng dụng có thể dễ dàng tích hợp các tính năng bổ sung:
  - **CORS** (Cross-Origin Resource Sharing) để cho phép truy cập từ client tại các domain khác nhau.
  - **Body-parser** hoặc tính năng tích hợp sẵn của Express để phân tích payload JSON, form data, file upload.
  - **Helmet** để thiết lập các header bảo mật cơ bản, giúp chống một số tấn công phổ biến (clickjacking, XSS, MIME sniffing...).
- Việc thêm mới hoặc thay thế middleware, route cũng rất đơn giản, giúp ứng dụng có thể mở rộng chức năng (ví dụ: logging, rate limiting, caching) mà không làm ảnh hưởng đến cấu trúc hiện tại.

- **Tích hợp dễ dàng với các hệ quản trị cơ sở dữ liệu:**

- Express.js không bó buộc vào một cơ sở dữ liệu cụ thể, có thể sử dụng song song với MongoDB (qua Mongoose), MySQL/PostgreSQL (qua Sequelize hoặc Knex.js), hoặc bất kỳ hệ quản trị nào khác. Trong đề tài, tùy theo đặc thù lưu trữ (từ vựng, lịch sử dịch, tiến độ ôn tập), nhóm có thể lựa chọn MongoDB để lưu trữ linh hoạt dạng document, hoặc MySQL/PostgreSQL nếu cần quan hệ chặt chẽ giữa các bảng.
- Nhờ tính mô-đun, phần data access layer (DAL) cũng được tách riêng, giúp dễ dàng thay đổi hoặc nâng cấp cơ sở dữ liệu mà không phải viết lại toàn bộ logic server.

- **Hiệu năng cao và cộng đồng hỗ trợ lớn:**

- Vì chạy trên Node.js, Express.js thừa hưởng cơ chế non-blocking

I/O, phù hợp để xử lý nhiều yêu cầu đồng thời, đặc biệt khi gọi các dịch vụ AI/ML (ví dụ API dịch AI, OCR) hay thực hiện các thao tác I/O vào cơ sở dữ liệu.

- Hệ sinh thái npm phong phú cho phép bổ sung hàng loạt thư viện hỗ trợ (ví dụ: multer cho upload hình ảnh, jsonwebtoken cho xác thực token, express-validator cho việc validate input...).
- Cộng đồng Node.js/Express.js lớn mạnh, nhiều plugin và tài liệu tham khảo, giúp quá trình xây dựng và gỡ lỗi trở nên thuận tiện hơn.

- **Hỗ trợ triển khai và mở rộng trên đám mây:**

- Express.js tương thích tốt với các dịch vụ đám mây như Heroku, AWS Elastic Beanstalk, Google Cloud Run... Việc đóng gói ứng dụng trong Docker hoặc tích hợp CI/CD pipeline cũng rất dễ dàng nhờ cấu trúc gọn gàng của ứng dụng Express.
- Nhờ đó, khi ứng dụng từ điển tiếng Nhật phát triển lớn hơn, cần tách microservices (chẳng hạn service riêng cho OCR, service riêng cho quản lý từ vựng), Express.js vẫn có thể đảm nhận vai trò gateway hoặc từng microservice độc lập.

Tóm lại, Express.js là lựa chọn lý tưởng cho server backend trong dự án của em vì tính nhẹ nhàng, dễ cấu hình và khả năng mở rộng cao. Với Express.js, việc xây dựng các endpoint phục vụ tra cứu, dịch câu, OCR và quản lý thẻ học được triển khai nhanh chóng, giúp tập trung phát triển các logic AI/FSRS cốt lõi mà không phải lo lắng nhiều về cấu hình phức tạp hay giới hạn framework.

### 2.3. Tổng quan về FastAPI



Hình 5 FastAPI

FastAPI là một framework web hiện đại, nhanh và hiệu quả dành cho Python, được thiết kế để xây dựng các API RESTful với hiệu suất cao, tương đương với Node.js và Go. FastAPI được xây dựng trên nền tảng của **Starlette** (xử lý web) và **Pydantic** (xử lý dữ liệu), hỗ trợ mạnh mẽ tính năng **type hints** của Python 3.6+ nhằm tự động kiểm tra, xác thực dữ liệu và sinh tài liệu API một cách trực quan.

Trong dự án xây dựng ứng dụng từ điển tiếng Nhật thông minh, FastAPI được lựa chọn để xây dựng backend vì sự gọn nhẹ, dễ phát triển, dễ bảo trì, đồng thời mang lại hiệu năng cao và khả năng mở rộng tốt cho các ứng dụng web hiện đại.

### Ưu điểm của FastAPI

- **Hiệu năng cao:** FastAPI được tối ưu hóa với uvicorn (ASGI server) và có hiệu suất xử lý yêu cầu tương đương với các framework tốc độ cao như Node.js hay Go.
- **Tự động sinh tài liệu API:** Với sự hỗ trợ của OpenAPI và Swagger UI, FastAPI tự động sinh tài liệu RESTful API trực quan tại các endpoint như /docs và /redoc. Điều này giúp lập trình viên dễ dàng kiểm thử và kiểm soát API.
- **Tích hợp Pydantic để kiểm tra dữ liệu:** FastAPI sử dụng Pydantic để xác thực đầu vào và đảm bảo dữ liệu đúng định dạng. Điều này làm giảm lỗi runtime và tăng độ an toàn khi xử lý dữ liệu.
- **Hỗ trợ async/await:** Framework hỗ trợ đầy đủ lập trình bất đồng bộ (asynchronous) giúp tối ưu hóa hiệu suất và khả năng xử lý đồng thời nhiều yêu cầu một cách hiệu quả.
- **Type hints rõ ràng:** Sử dụng type hints của Python giúp viết mã rõ ràng, dễ hiểu và tận dụng tốt IDE như VSCode để gợi ý và kiểm tra lỗi.
- **Dễ triển khai microservices:** Nhờ sự đơn giản và hiệu quả, FastAPI rất phù hợp để xây dựng các hệ thống microservices nhỏ gọn hoặc làm backend cho các frontend SPA (Single Page Applications).

### Ứng dụng trong đề tài

Trong đề tài này, FastAPI đóng vai trò là backend cho các tác vụ load model cho ứng dụng. Một số chức năng chính sử dụng FastAPI gồm:

- Nhận diện chữ kanji trên mỗi nét chữ mà người dùng vẽ giúp người dùng tra cứu chữ kanji nhanh chóng.
- Detect giọng nói của người dùng và chuyển thành văn bản.
- Sử dụng model dịch thuật và trả về bản dịch realtime, chính xác cho người dùng.

## 2.4. Tổng quan về MongoDB



Hình 6 Logo MongoDB

MongoDB là hệ quản trị cơ sở dữ liệu NoSQL mã nguồn mở, lưu trữ dữ liệu theo dạng tài liệu (document) thay vì bảng (table) như hệ quản trị quan hệ (RDBMS). Dữ liệu trong MongoDB được lưu dưới định dạng BSON (Binary JSON), giúp hỗ trợ linh hoạt cấu trúc dữ liệu và mở rộng dễ dàng. Thay vì cố định các bảng và cột, MongoDB sử dụng các collection (tập hợp các document), trong đó mỗi document là một đối tượng JSON có thể chứa các trường (field) đa dạng, bao gồm các mảng (array) và đối tượng lồng nhau (embedded document).

MongoDB hoạt động theo kiến trúc máy chủ—máy khách tương tự như các hệ quản trị khác. Máy chủ MongoDB (mongod) chịu trách nhiệm lưu trữ, xử lý truy vấn và quản lý dữ liệu. Các ứng dụng hoặc công cụ quản trị (mongo shell, MongoDB Compass, hoặc driver ngôn ngữ) kết nối tới mongod qua giao thức mạng để thực hiện các thao tác CRUD (Create, Read, Update, Delete). Máy chủ MongoDB có thể triển khai trên một máy chủ độc lập hoặc trong cụm cluster, hỗ trợ replica set (sao chép dữ liệu) và sharding (phân mảnh) để tăng tính sẵn sàng và khả năng mở rộng.

#### **Ưu điểm của MongoDB:**

- **Mô hình dữ liệu linh hoạt (Schema-less):** Mỗi document có thể có cấu trúc khác nhau, không cần định nghĩa sẵn schema. Điều này cho phép ứng dụng dễ thay đổi, mở rộng nhanh chóng khi yêu cầu dữ liệu phát triển theo thời gian.
- **Hiệu suất cao với dữ liệu phi cấu trúc:** Vì dữ liệu được lưu dưới dạng BSON, truy vấn document thường nhanh và giảm thiểu việc join phức tạp. MongoDB tận dụng bộ nhớ RAM để cache dữ liệu nóng (working set), giúp tăng tốc độ đọc/ghi.
- **Dễ dàng mở rộng ngang (Horizontal Scalability):** Với sharding, dữ liệu được phân tán (partition) trên nhiều shard, mỗi shard là một replica set. Architecture này giúp xử lý lượng lớn dữ liệu và lưu lượng truy cập mà không làm giảm hiệu năng.
- **Hỗ trợ các tính năng NoSQL tiên tiến:**
  - **Replica Set:** Tự động sao chép dữ liệu giữa các node, đảm bảo tính sẵn sàng và dự phòng khi có sự cố.

- **Sharding:** Phân mảnh dữ liệu theo key chuyên biệt, cho phép mở rộng cluster linh hoạt khi khối lượng dữ liệu tăng.
  - **Aggregation Framework:** Bộ công cụ mạnh mẽ để xử lý, biến đổi và tổng hợp dữ liệu ngay trong cơ sở dữ liệu.
  - **Full-Text Search & Geospatial Queries:** Hỗ trợ tìm kiếm văn bản đầy đủ và truy vấn không gian địa lý, phù hợp với nhiều ứng dụng hiện đại.
- **Hệ sinh thái phong phú:** MongoDB cung cấp các driver chính thức cho nhiều ngôn ngữ (Node.js, Java, Python, Go, v.v.), các công cụ quản trị như MongoDB Compass, Atlas (dịch vụ quản lý cloud), giúp việc phát triển và triển khai trở nên đơn giản.
  - **Triển khai trên đám mây và on-premises:** Người dùng có thể cài đặt MongoDB trên server riêng, hoặc sử dụng MongoDB Atlas – dịch vụ DBaaS trên AWS, Google Cloud, Azure—để tận hưởng tính năng tự động sao lưu, update và giám sát.

### Nhược điểm của MongoDB

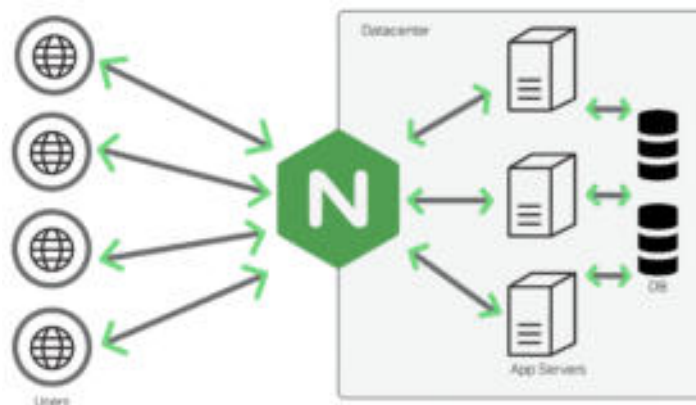
- **Không có schema ràng buộc (Schema-less) có thể gây rối:** Dù linh hoạt, việc không định nghĩa schema cứng nhắc cũng dễ dẫn đến dữ liệu không đồng nhất, khó khăn trong việc kiểm soát và validate. Cần áp dụng validation tại cấp độ ứng dụng hoặc schema validation của MongoDB để đảm bảo tính nhất quán.
- **Tính nhất quán dữ liệu (Consistency):** Trong mô hình phân tán, MongoDB ưu tiên tính sẵn sàng và phân vùng (availability & partition tolerance), đôi khi phải đánh đổi tính nhất quán mạnh (strong consistency). Khi thực hiện ghi chép trên replica set, có độ trễ giữa các node, dẫn đến trường hợp đọc dữ liệu “lỗi thời” (eventual consistency).
- **Hạn chế trong giao dịch (Transaction):** Phiên bản trước MongoDB 4.0 không hỗ trợ giao dịch đa document (multi-document transaction). Kể từ MongoDB 4.0+, đã bổ sung khả năng transaction ACID xuyên collection, nhưng hiệu suất có thể kém hơn so với RDBMS truyền thống khi xử lý các giao dịch phức tạp.
- **Quản lý bộ nhớ và ổ đĩa:** Vì lưu trữ dưới dạng BSON (Binary JSON), kích thước dữ liệu thường lớn hơn so với định dạng nhị phân riêng của RDBMS. Cần cân nhắc cấu hình bộ nhớ RAM đủ lớn để giữ working set, tránh trường hợp thrashing (đổi trang liên tục) làm giảm hiệu năng.
- **Chi phí vận hành có thể cao:** Khi mở rộng cluster sharded với nhiều replica set, sẽ phát sinh chi phí hạ tầng đáng kể. Nếu sử dụng MongoDB Atlas, chi phí dịch vụ

cloud cũng cần tính toán cho phù hợp với nhu cầu lưu trữ và băng thông.

Tóm lại, với lợi thế về mô hình document-oriented, khả năng mở rộng dễ dàng và tích hợp tốt cùng Node.js, MongoDB là giải pháp cơ sở dữ liệu phù hợp cho ứng dụng từ điển tiếng Nhật thông minh, đáp ứng yêu cầu linh hoạt, hiệu năng cao và sẵn sàng phát triển trong tương lai.

## 2.5. Tổng quan về Nginx

Nginx là một web server mạnh mẽ mã nguồn mở. Nginx sử dụng kiến trúc đơn luồng, hướng sự kiện vì thế nó hiệu quả hơn Apache server. Nó cũng có thể làm những thứ quan trọng khác, chẳng hạn như load balancing, HTTP caching, hay sử dụng như một reverse proxy. Nginx là kiến thức không thể thiếu đối với một web developer, system administrator hay devops.



Hình 7 Nginx

Những tính năng của máy chủ HTTP Nginx:

- Có khả năng xử lý hơn 10.000 kết nối cùng lúc với bộ nhớ thấp;
- Phục vụ tập tin tĩnh (static files) và lập chỉ mục tập tin;
- Tăng tốc reverse proxy bằng bộ nhớ đệm (cache), cân bằng tải đơn giản và khả năng chịu lỗi;
- Hỗ trợ tăng tốc với bộ nhớ đệm của FastCGI, uWSGI, SCGI và các máy chủ memcached;
- Hỗ trợ mã hóa SSL và TLS;
- Cấu hình của Nginx vô cùng linh hoạt giúp lưu lại nhật ký truy vấn một cách dằn;
- Nginx có khả năng chuyển hướng lỗi 3XX-5XX;
- Rewrite URL có thể sử dụng expression;
- Nginx giúp giới hạn số kết nối đồng thời cũng như truy vấn từ 1 địa chỉ;
- Nginx có khả năng nhúng mã PERL một cách dễ dàng;

- Nginx có thể hỗ trợ và tương thích hoàn toàn với IPv6;
- Nginx có thể hỗ trợ cho websockets;
- Nginx hỗ trợ truyền tải các file FLV và MP4.

## 2.6. Tổng quan về Docker



Hình 8 Docker

Docker là một nền tảng mã nguồn mở dùng để phát triển, vận chuyển và chạy các ứng dụng. Docker cho phép bạn tách ứng dụng ra khỏi cơ sở hạ tầng, giúp việc phân phối phần mềm trở nên nhanh chóng và linh hoạt hơn. Với Docker, bạn có thể quản lý cơ sở hạ tầng tương tự như cách bạn quản lý mã nguồn của ứng dụng. Bằng cách tận dụng các phương pháp vận chuyển, kiểm thử và triển khai mã mà Docker cung cấp, bạn có thể giảm đáng kể độ trễ giữa giai đoạn phát triển và đưa mã vào môi trường sản xuất. Container (hay vùng chứa) cung cấp môi trường nhẹ, nhất quán bao gồm ứng dụng và toàn bộ các phần phụ thuộc cần thiết để chạy ứng dụng đó.

Đặc trưng cơ bản của Docker:

- Consistency - Tính nhất quán: Loại bỏ các vấn đề kiểu "nó chạy được trên máy của tôi" bằng cách đảm bảo môi trường phát triển đồng nhất. Tất cả lập trình viên đều sử dụng cùng một môi trường, giúp dự án chạy ổn định và dễ tái tạo.
- Isolation - Tính cô lập, độc lập: Mỗi container hoạt động độc lập với các ứng dụng khác, giúp ngăn ngừa xung đột phần mềm và đảm bảo môi trường riêng biệt cho từng ứng dụng.
- Portability - Tính di động, linh động: Các container có thể chạy trên bất kỳ hệ thống nào hỗ trợ Docker, bất kể nền tảng cơ sở hạ tầng bên dưới là gì (Windows, macOS, Linux, máy thật hay máy ảo).
- Efficiency - Tính hiệu quả: Các container chia sẻ nhân hệ điều hành máy chủ, giảm chi phí tài nguyên.

## 2.7. Dịch vụ EC2 trên điện toán đám mây AWS



Hình 9 Elastic Compute Cloud – Amazon EC2

Elastic Compute Cloud – Amazon EC2 là một dịch vụ cung cấp máy chủ ảo trên nền tảng điện toán đám mây do Amazon Web Services phát triển. EC2 cho phép người dùng dễ dàng tạo, cấu hình và vận hành các máy chủ ảo (instances) với nhiều hệ điều hành, cấu hình phần cứng và phần mềm khác nhau tùy theo nhu cầu sử dụng.

Các tính năng chính của EC2:

- **Tùy chỉnh linh hoạt:** Người dùng có thể lựa chọn loại máy chủ phù hợp với nhu cầu về CPU, bộ nhớ, dung lượng lưu trữ, và mạng.
- **Khả năng mở rộng:** EC2 hỗ trợ auto-scaling, giúp tự động tăng hoặc giảm số lượng máy chủ khi lưu lượng thay đổi.
- **Bảo mật:** EC2 hỗ trợ cấu hình firewall thông qua Security Groups và các tính năng bảo mật mạng khác như Virtual Private Cloud (VPC).
- **Tích hợp mạnh mẽ:** EC2 có thể tích hợp với các dịch vụ khác của AWS như S3, RDS, CloudWatch, IAM,... để xây dựng hệ thống hoàn chỉnh.
- **Thanh toán linh hoạt:** Người dùng chỉ phải trả phí cho thời gian thực tế sử dụng tài nguyên.

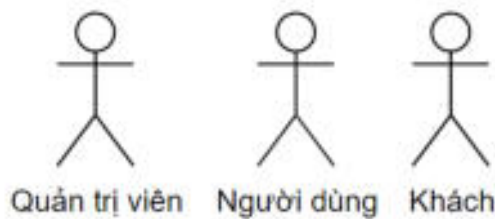
Việc sử dụng EC2 giúp đảm bảo rằng hệ thống có hiệu suất ổn định, dễ mở rộng và có thể tùy chỉnh linh hoạt để phục vụ số lượng người dùng lớn mà vẫn đảm bảo tính bảo mật và hiệu quả vận hành.

## CHƯƠNG 3. PHÂN TÍCH THIẾT KẾ HỆ THỐNG

### 3.1. Các tác nhân chính của hệ thống

Trang web được xây dựng bao gồm 3 tác nhân chính hoạt động:

- Quản trị viên
- Người dùng
- Khách



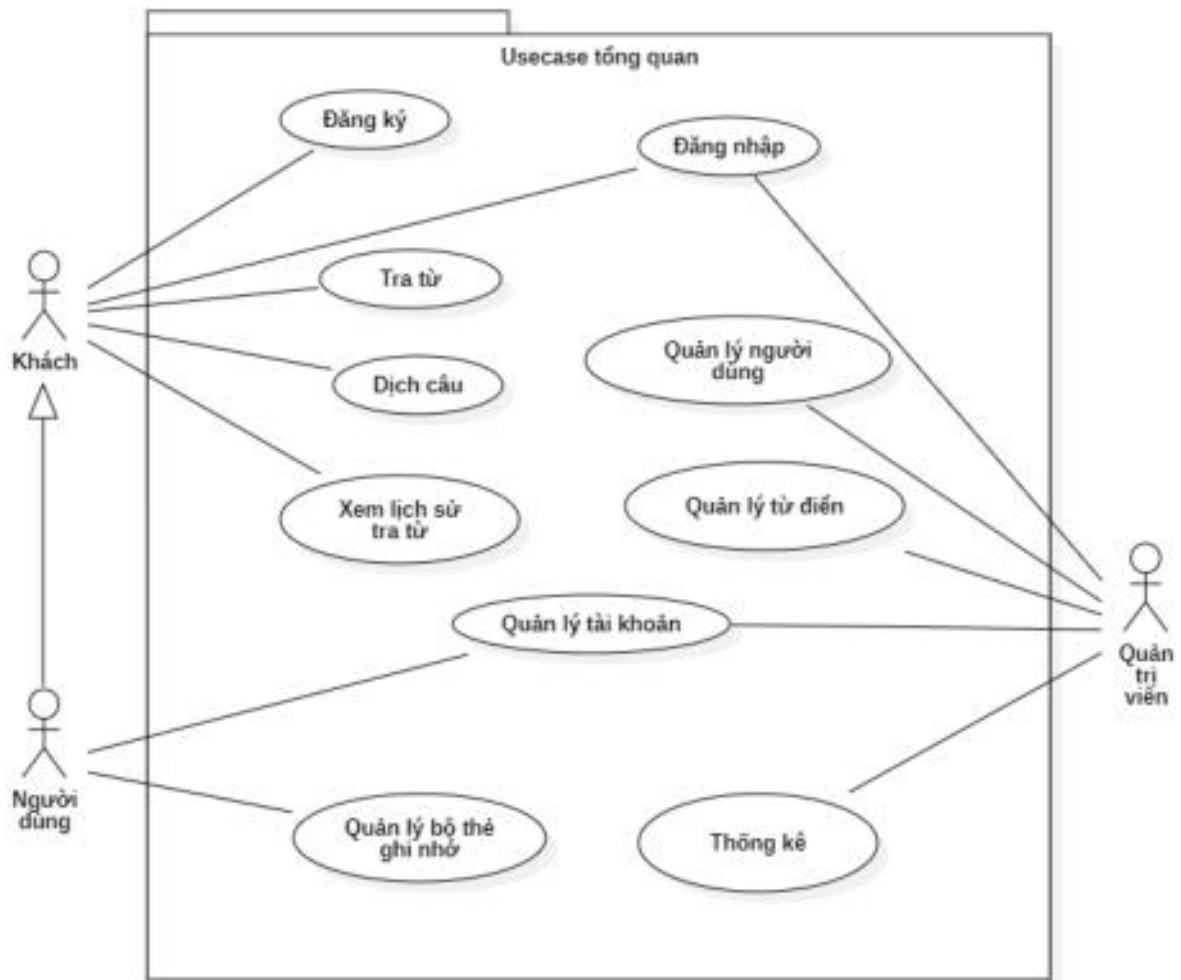
Hình 10 Các tác nhân trong hệ thống

| Tên tác nhân  | Chức năng chính                                                                                                                                                                                                                                                                                                                                                |
|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Quản trị viên | <ul style="list-style-type: none"> <li>- Thêm, sửa xoá từ vựng và hán tự</li> <li>- Xem danh sách người dùng, chi tiết người dùng.</li> <li>- Xem biểu đồ dữ liệu người dùng, từ vựng và hán tự</li> </ul>                                                                                                                                                     |
| Người dùng    | <ul style="list-style-type: none"> <li>- Quản lý tài khoản</li> <li>- Quản lý được bộ thẻ và thẻ ghi nhớ</li> <li>- Học từ vựng, hán tự</li> <li>- Xem bảng thống kê học tập</li> <li>- Tìm kiếm từ vựng hán tự</li> <li>- Xem lịch sử tìm kiếm</li> <li>- Dịch từ tiếng Nhật sang Việt hoặc Việt sang Nhật từ câu văn bản ngắn hoặc là từ hình ảnh</li> </ul> |
| Khách         | <ul style="list-style-type: none"> <li>- Tìm kiếm từ vựng hán tự</li> <li>- Xem lịch sử tìm kiếm</li> <li>- Dịch từ tiếng Nhật sang Việt hoặc Việt sang Nhật từ câu văn bản ngắn hoặc là từ hình ảnh</li> </ul>                                                                                                                                                |

Bảng 1 Các chức năng chính của hệ thống

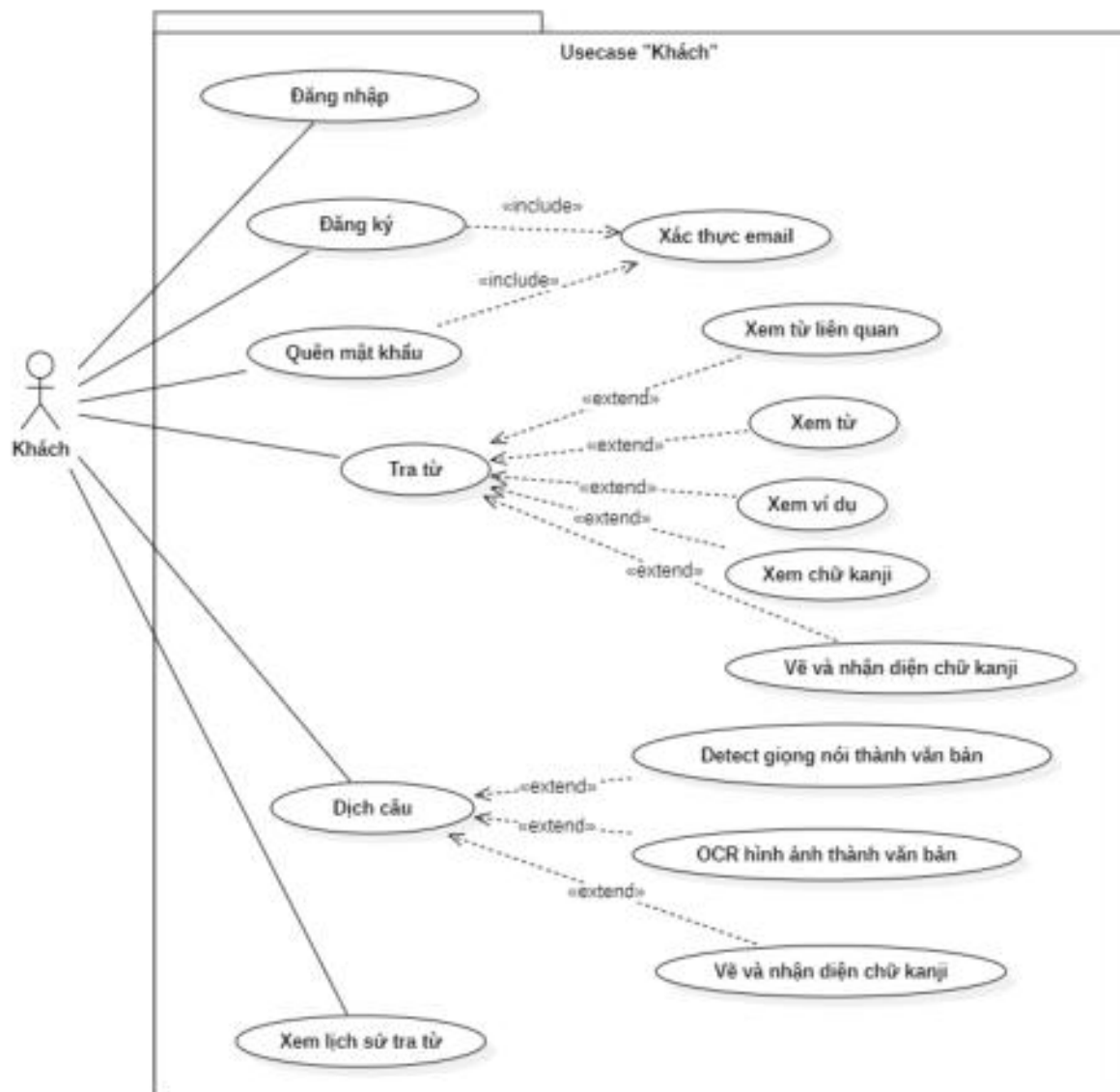
### 3.2. Biểu đồ usecase

#### 3.2.1. Biểu đồ usecase tổng quan

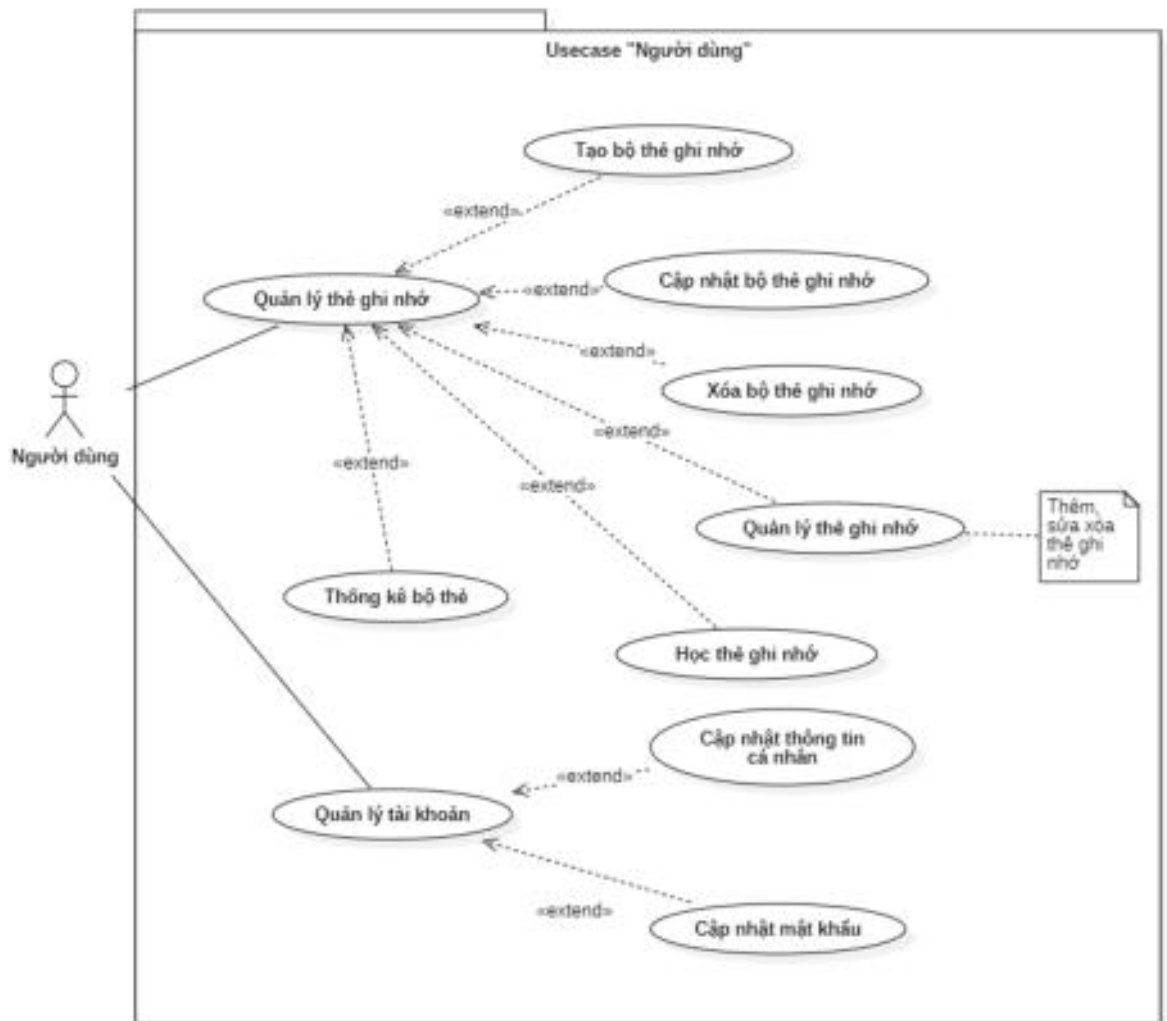


Hình 11 Biểu đồ usecase tổng quan

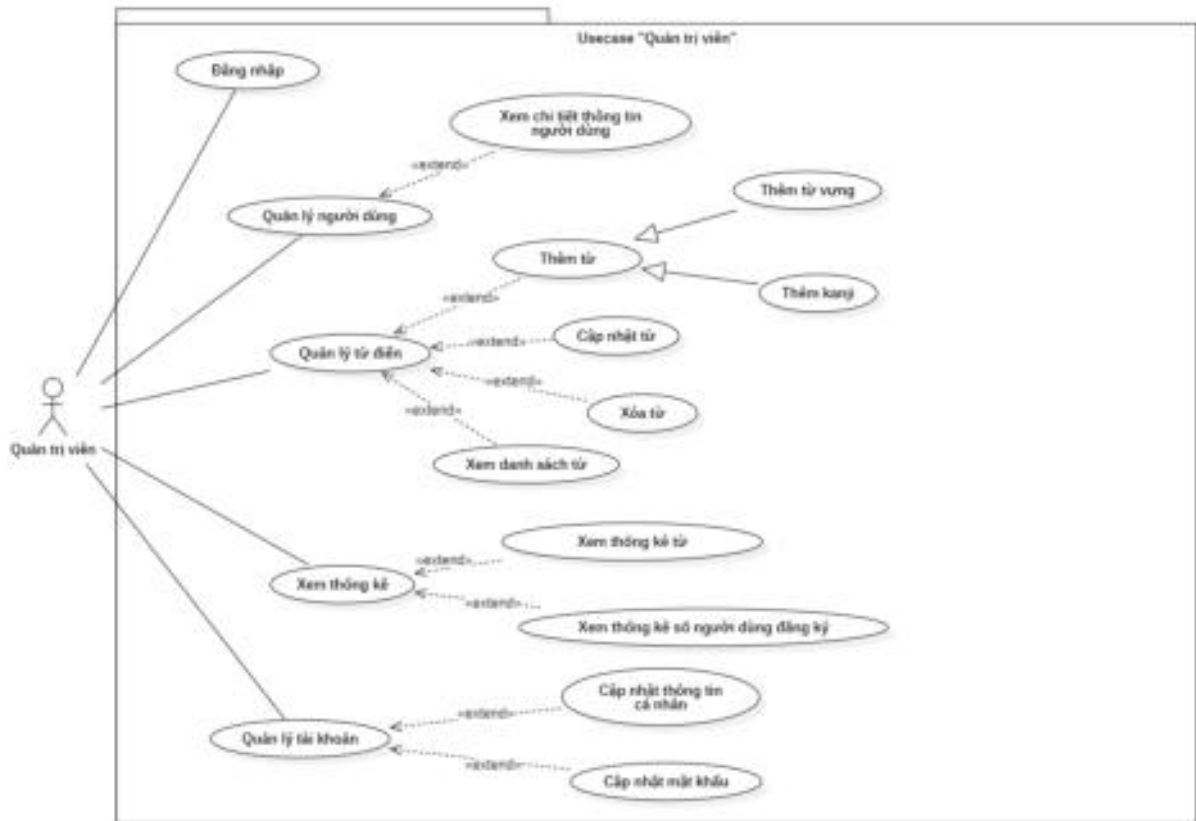
### 3.2.2. Biểu đồ usecase phân rã



Hình 12 Biểu đồ usecase phân rã theo “Khách”



Hình 13 Biểu đồ usecase phân rã theo “Người dùng”



Hình 14 Biểu đồ usecase phân rã theo “Quản trị viên”

### 3.3. Đặc tả biểu đồ usecase

#### 3.3.1. Tra từ

|                                                                                                               |                                               |
|---------------------------------------------------------------------------------------------------------------|-----------------------------------------------|
| <b>Mã Usecase:</b> UC-01                                                                                      |                                               |
| <b>Tên Usecase:</b> Tra từ                                                                                    |                                               |
| <b>Mô tả:</b> Tra cứu từ vựng hoặc hán tự                                                                     |                                               |
| <b>Tác nhân:</b> Khách, người dùng                                                                            |                                               |
| <b>Tiền điều kiện:</b> Chưa hoặc đã đăng nhập vào hệ thống, truy cập vào giao diện chính và ấn thanh tìm kiếm |                                               |
| <b>Hậu điều kiện:</b> Hệ thống hiển thị thông tin kết quả tìm kiếm tương ứng                                  |                                               |
| <b>Luồng sự kiện chính</b>                                                                                    |                                               |
| <b>Tác nhân</b>                                                                                               | <b>Hệ thống</b>                               |
| 1.Chọn giao diện chính                                                                                        |                                               |
| 2.Chọn vào thanh tìm kiếm                                                                                     |                                               |
| 3.Nhập từ muốn tìm kiếm                                                                                       |                                               |
|                                                                                                               | 4.Hiển thị danh sách kết quả tìm kiếm         |
| <b>Luồng sự kiện thay thế</b>                                                                                 |                                               |
| 3.1.Vẽ chữ kanji muốn tìm kiếm                                                                                |                                               |
| 3.2. Ghi âm từ/câu muốn tìm kiếm                                                                              |                                               |
|                                                                                                               | 4.1.Hiển thị thông báo không tìm thấy kết quả |

Bảng 2 Đặc tả usecase tra từ

### 3.3.2. Dịch câu

|                                                                                                                                     |                        |
|-------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| <b>Mã Usecase:</b> UC-02                                                                                                            |                        |
| <b>Tên Usecase:</b> Dịch câu                                                                                                        |                        |
| <b>Mô tả:</b> Dịch đoạn văn bản ngắn, văn bản ngắn trong hình ảnh, văn bản ngắn từ giọng nói ghi âm, từ tiếng Việt sang tiếng Nhật. |                        |
| <b>Tác nhân:</b> Khách, người dùng                                                                                                  |                        |
| <b>Tiền điều kiện:</b> Chưa hoặc đã đăng nhập vào hệ thống, truy cập vào giao diện dịch                                             |                        |
| <b>Hậu điều kiện:</b> Hệ thống trả về kết quả dịch của câu                                                                          |                        |
| <b>Luồng sự kiện chính</b>                                                                                                          |                        |
| <b>Tác nhân</b>                                                                                                                     | <b>Hệ thống</b>        |
| 1.Chọn vào giao diện dịch                                                                                                           |                        |
| 2.Chọn ngôn ngữ Nhật dịch sang Việt                                                                                                 |                        |
| 3.Nhập đoạn văn bản ngắn cần dịch                                                                                                   |                        |
| 4.Ấn vào nút “Dịch”                                                                                                                 |                        |
|                                                                                                                                     | 5.Sinh ra văn bản dịch |
| <b>Luồng sự kiện thay thế</b>                                                                                                       |                        |
| 2.1.Chọn ngôn ngữ Việt sang Nhật                                                                                                    |                        |
| 3.1.Chọn hình ảnh chứa văn bản cần dịch                                                                                             |                        |
| 3.2. Ghi âm giọng nói ra văn bản cần dịch                                                                                           |                        |

Bảng 3 Đặc tả usecase dịch câu

### 3.3.3. Xem lịch sử tra từ

|                                                                                                                                                       |                                |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|
| <b>Mã Usecase:</b> UC-03                                                                                                                              |                                |
| <b>Tên Usecase:</b> Xem lịch sử tra từ                                                                                                                |                                |
| <b>Mô tả:</b> Xem lịch sử tra từ theo hán tự và từ vựng                                                                                               |                                |
| <b>Tác nhân:</b> Khách hàng, người dùng                                                                                                               |                                |
| <b>Tiền điều kiện:</b> Chưa hoặc đã đăng nhập vào hệ thống, truy cập vào trang giao diện chính hoặc trang giao diện dịch hoặc nhấn vào thanh tìm kiếm |                                |
| <b>Hậu điều kiện:</b> Hiện thị danh sách lịch sử                                                                                                      |                                |
| <b>Luồng sự kiện chính</b>                                                                                                                            |                                |
| <b>Tác nhân</b>                                                                                                                                       | <b>Hệ thống</b>                |
| 1.Chọn giao diện chính hoặc giao diện dịch hoặc nhấn vào thanh tìm kiếm.                                                                              |                                |
|                                                                                                                                                       | 2.Hiện thị ra lịch sử tìm kiếm |

Bảng 4 Đặc tả usecase xem lịch sử tra từ

### 3.3.4. Đăng nhập

|                                                     |                                                                                                 |
|-----------------------------------------------------|-------------------------------------------------------------------------------------------------|
| <b>Mã Usecase:</b> UC-4                             |                                                                                                 |
| <b>Tên Usecase:</b> Đăng nhập                       |                                                                                                 |
| <b>Mô tả:</b> Đăng nhập vào hệ thống                |                                                                                                 |
| <b>Tác nhân:</b> Khách, Quản trị viên               |                                                                                                 |
| <b>Tiền điều kiện:</b> Đã có tài khoản của hệ thống |                                                                                                 |
| <b>Hậu điều kiện:</b> Đăng nhập vào hệ thống        |                                                                                                 |
| <b>Luồng sự kiện chính</b>                          |                                                                                                 |
| <b>Tác nhân</b>                                     | <b>Hệ thống</b>                                                                                 |
| 1.Chọn vào nút “Đăng nhập”                          |                                                                                                 |
| 2.Nhập thông tin email                              |                                                                                                 |
| 3. Nhập thông tin mật khẩu                          |                                                                                                 |
| 4.Chọn nút “Đăng nhập”                              |                                                                                                 |
|                                                     | 5.Hệ thống hiển thị thông báo “Đăng nhập thành công” và hiển thị giao diện                      |
| <b>Luồng sự kiện thay thế</b>                       |                                                                                                 |
|                                                     | 5.1.Hệ thống hiển thị thông báo”Email hoặc mật khẩu không đúng”                                 |
|                                                     | 5.2.Hệ thống hiển thị thông báo”Một email đã được gửi đến tài khoản của bạn, vui lòng xác thực” |

Bảng 5 Đặc tả usecase đăng nhập

### 3.3.5. Đăng ký

|                                                                            |                                                                                               |
|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|
| <b>Mã Usecase:</b> UC-5                                                    |                                                                                               |
| <b>Tên Usecase:</b> Đăng ký                                                |                                                                                               |
| <b>Mô tả:</b> Đăng ký                                                      |                                                                                               |
| <b>Tác nhân:</b> Khách                                                     |                                                                                               |
| <b>Tiền điều kiện:</b> Chưa có tài khoản của hệ thống.                     |                                                                                               |
| <b>Hậu điều kiện:</b> Thêm thông tin tài khoản của người dùng vào hệ thống |                                                                                               |
| <b>Luồng sự kiện chính</b>                                                 |                                                                                               |
| <b>Tác nhân</b>                                                            | <b>Hệ thống</b>                                                                               |
| 1.Chọn vào nút “Đăng ký”                                                   |                                                                                               |
| 2.Nhập email                                                               |                                                                                               |
| 3.Nhập mật khẩu                                                            |                                                                                               |
| 4.Nhập xác nhận mật khẩu                                                   |                                                                                               |
| 5.Nhấn vào nút “Đăng ký”                                                   |                                                                                               |
|                                                                            | 6.Hệ thống hiển thị thông báo thành công, hệ thống gửi đường link xác thực vào email đăng ký. |
| 7.Nhấn vào đường link xác thực để xác thực tài khoản.                      |                                                                                               |
|                                                                            | 8.Hệ thống hiển thị giao diện trang chủ                                                       |
| <b>Luồng sự kiện thay thế</b>                                              |                                                                                               |
|                                                                            | 6.1.Hệ thống hiển thị thông báo “Email                                                        |

|  |                                               |
|--|-----------------------------------------------|
|  | đã tồn tại, vui lòng đăng ký với email khác!” |
|--|-----------------------------------------------|

Bảng 6. Đặc tả usecase đăng ký

## 3.3.6. Quên mật khẩu

|                                                                                      |                                                                                            |
|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| <b>Mã Usecase:</b> UC-06                                                             |                                                                                            |
| <b>Tên Usecase:</b> Quên mật khẩu                                                    |                                                                                            |
| <b>Mô tả:</b> Quên mật khẩu                                                          |                                                                                            |
| <b>Tác nhân:</b> Khách hàng                                                          |                                                                                            |
| <b>Tiền điều kiện:</b> Có tài khoản của hệ thống, truy cập vào giao diện “đăng nhập” |                                                                                            |
| <b>Hậu điều kiện:</b> Cập nhật thông tin tài khoản của người dùng vào hệ thống       |                                                                                            |
| <b>Luồng sự kiện chính</b>                                                           |                                                                                            |
| <b>Tác nhân</b>                                                                      | <b>Hệ thống</b>                                                                            |
| 1.Chọn vào “Đăng nhập”                                                               |                                                                                            |
| 2.Chọn vào “Quên mật khẩu”                                                           |                                                                                            |
| 3.Nhập email                                                                         |                                                                                            |
| 4.Chọn gửi về email                                                                  |                                                                                            |
|                                                                                      | 5.Hệ thống gửi link xác thực về email                                                      |
| 5.Người dùng nhấn vào đường link xác thực                                            |                                                                                            |
|                                                                                      | 6.Hệ thống hiển thị giao diện “Nhập mật khẩu mới.”                                         |
| 7.Nhập mật khẩu mới                                                                  |                                                                                            |
| 8.Nhập xác nhận mật khẩu                                                             |                                                                                            |
| 9. Nhấn nút “Thay đổi mật khẩu”                                                      |                                                                                            |
|                                                                                      | 10.Hệ thống thông báo “Đổi mật khẩu thành công”                                            |
| <b>Luồng sự kiện thay thế</b>                                                        |                                                                                            |
|                                                                                      | 6.1.Hệ thống thông báo “OTP sai”                                                           |
|                                                                                      | 10.1.Hệ thống thông báo “Mật khẩu và mật khẩu xác nhận không chính xác”                    |
|                                                                                      | 10.2. Hệ thống thông báo “Mật khẩu này đã được dùng trước đó. Vui long nhập mật khẩu mới.” |

Bảng 7. Đặc tả usecase quên mật khẩu

## 3.3.7. Tạo bộ thẻ ghi nhớ

|                                                                             |  |
|-----------------------------------------------------------------------------|--|
| <b>Mã Usecase:</b> UC-07                                                    |  |
| <b>Tên Usecase:</b> Tạo bộ thẻ ghi nhớ                                      |  |
| <b>Mô tả:</b> Tạo bộ thẻ ghi nhớ                                            |  |
| <b>Tác nhân:</b> Người dùng                                                 |  |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống, truy cập giao diện bộ thẻ |  |
| <b>Hậu điều kiện:</b> Bộ thẻ được thêm vào hệ thống                         |  |
| <b>Luồng sự kiện chính</b>                                                  |  |

| Tác nhân                      | Hệ thống                    |
|-------------------------------|-----------------------------|
| 1.Nhấn vào nút “Add”          |                             |
| 2.Nhập vào tên bộ thẻ         |                             |
| 3.Nhấn nút “Tạo”              |                             |
|                               | 6.Hệ thống thêm bộ thẻ mới. |
| <b>Luồng sự kiện thay thế</b> |                             |
|                               |                             |

Bảng 8 Đặc tả usecase tạo bộ thẻ ghi nhớ

## 3.3.8. Chỉnh sửa bộ thẻ ghi nhớ

| <b>Mã Usecase:</b> UC-08                                                    |                                                            |
|-----------------------------------------------------------------------------|------------------------------------------------------------|
| <b>Tên Usecase:</b> Chỉnh sửa bộ thẻ ghi nhớ                                |                                                            |
| <b>Mô tả:</b> Chỉnh sửa bộ thẻ ghi nhớ                                      |                                                            |
| <b>Tác nhân:</b> Người dùng                                                 |                                                            |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống, truy cập giao diện bộ thẻ |                                                            |
| <b>Hậu điều kiện:</b> Thông tin bộ thẻ mới được cập nhật trong hệ thống     |                                                            |
| <b>Luồng sự kiện chính</b>                                                  |                                                            |
| Tác nhân                                                                    | Tác nhân                                                   |
| 1.Nhấn vào bộ thẻ muốn chỉnh sửa (Nhấn vào biểu tượng bánh răng)            |                                                            |
| 2.Nhấn vào nút “Rename”                                                     |                                                            |
| 3.Nhập vào tên bộ thẻ mới                                                   |                                                            |
| 4.Nhấn nút “Lưu”                                                            |                                                            |
|                                                                             | 5.Hệ thống cập nhật thông tin mới của bộ thẻ vào giao diện |
| <b>Luồng sự kiện thay thế</b>                                               |                                                            |
|                                                                             |                                                            |

Bảng 9 Đặc tả usecase chỉnh sửa bộ thẻ ghi nhớ

## 3.3.9. Xóa bộ thẻ ghi nhớ

| <b>Mã Usecase:</b> UC-9                                                     |                                  |
|-----------------------------------------------------------------------------|----------------------------------|
| <b>Tên Usecase:</b> Xóa bộ thẻ ghi nhớ                                      |                                  |
| <b>Mô tả:</b> Xóa bộ thẻ ghi nhớ                                            |                                  |
| <b>Tác nhân:</b> Người dùng                                                 |                                  |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống, truy cập giao diện bộ thẻ |                                  |
| <b>Hậu điều kiện:</b> Bộ thẻ ghi nhớ bị xóa khỏi hệ thống                   |                                  |
| <b>Luồng sự kiện chính</b>                                                  |                                  |
| Tác nhân                                                                    | Hệ thống                         |
| 1.Nhấn vào bộ thẻ muốn xóa (Nhấn vào biểu tượng bánh răng)                  |                                  |
| 2.Nhấn vào nút “Delete”                                                     |                                  |
|                                                                             | 3.Hệ thống xóa bộ thẻ tương ứng. |
| <b>Luồng sự kiện thay thế</b>                                               |                                  |
|                                                                             |                                  |

Bảng 10 Đặc tả usecase xóa bộ thẻ ghi nhớ

**3.3.10. Tạo thẻ ghi nhớ**

|                                                                             |                                    |
|-----------------------------------------------------------------------------|------------------------------------|
| <b>Mã Usecase:</b> UC-10                                                    |                                    |
| <b>Tên Usecase:</b> Tạo thẻ ghi nhớ                                         |                                    |
| <b>Mô tả:</b> Tạo thẻ ghi nhớ                                               |                                    |
| <b>Tác nhân:</b> Người dùng                                                 |                                    |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống, truy cập giao diện bộ thẻ |                                    |
| <b>Hậu điều kiện:</b> Thẻ mới được thêm vào bộ thẻ tương ứng                |                                    |
| <b>Luồng sự kiện chính</b>                                                  |                                    |
| <b>Tác nhân</b>                                                             | <b>Hệ thống</b>                    |
| 1.Nhấn vào biểu tượng bánh răng                                             |                                    |
| 2.Nhấn vào nút “Add card”                                                   |                                    |
| 3. Chọn bộ thẻ muốn thêm thẻ ghi nhớ                                        |                                    |
| 4.Nhập từ vựng                                                              |                                    |
| 5.Nhập câu ví dụ                                                            |                                    |
| 6.Nhập cách đọc                                                             |                                    |
| 7.Nhập ý nghĩa                                                              |                                    |
| 8.Nhấn vào “Save”                                                           |                                    |
|                                                                             | 9.Tạo thẻ mới vào bộ thẻ tương ứng |
| <b>Luồng sự kiện thay thế</b>                                               |                                    |
|                                                                             |                                    |

Bảng 11 Đặc tả usecase tạo thẻ ghi nhớ

**3.3.11. Sửa thẻ ghi nhớ**

|                                                                                     |                                       |
|-------------------------------------------------------------------------------------|---------------------------------------|
| <b>Mã Usecase:</b> UC-11                                                            |                                       |
| <b>Tên Usecase:</b> Chỉnh sửa thông tin thẻ ghi nhớ                                 |                                       |
| <b>Mô tả:</b> Chỉnh sửa thông tin thẻ ghi nhớ                                       |                                       |
| <b>Tác nhân:</b> Người dùng                                                         |                                       |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống, truy cập vào giao diện học bộ thẻ |                                       |
| <b>Hậu điều kiện:</b> Thẻ được cập nhật thông tin                                   |                                       |
| <b>Luồng sự kiện chính</b>                                                          |                                       |
| <b>Tác nhân</b>                                                                     | <b>Hệ thống</b>                       |
| 1.Nhấn vào icon “Chỉnh sửa” thẻ                                                     |                                       |
| 2.Nhập từ vựng                                                                      |                                       |
| 3.Nhập câu ví dụ                                                                    |                                       |
| 4.Nhập cách đọc                                                                     |                                       |
| 6.Nhập ý nghĩa                                                                      |                                       |
| 7.Nhấn vào nút “Lưu”                                                                |                                       |
|                                                                                     | 8.Hệ thống cập nhật thông tin của thẻ |
| <b>Luồng sự kiện thay thế</b>                                                       |                                       |
|                                                                                     |                                       |

Bảng 12 Đặc tả usecase sửa thẻ ghi nhớ

**3.3.12. Xóa thẻ ghi nhớ**

|                                     |  |
|-------------------------------------|--|
| <b>Mã Usecase:</b> UC-12            |  |
| <b>Tên Usecase:</b> Xóa thẻ ghi nhớ |  |

|                                                                                     |                                     |
|-------------------------------------------------------------------------------------|-------------------------------------|
| <b>Mô tả:</b> Xóa thẻ ghi nhớ                                                       |                                     |
| <b>Tác nhân:</b> Người dùng                                                         |                                     |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống, truy cập vào giao diện học bộ thẻ |                                     |
| <b>Hậu điều kiện:</b> Thẻ được xóa khỏi hệ thống                                    |                                     |
| <b>Luồng sự kiện chính</b>                                                          |                                     |
| <b>Tác nhân</b>                                                                     | <b>Hệ thống</b>                     |
| 1.Nhấn vào icon “Xóa” thẻ                                                           |                                     |
| 2.Nhấn vào “Xóa” ở modal xác nhận                                                   |                                     |
|                                                                                     | 3.Hệ thống xóa thẻ ra khỏi hệ thống |
| <b>Luồng sự kiện thay thế</b>                                                       |                                     |
|                                                                                     |                                     |

Bảng 13 Đặc tả usecase xóa thẻ ghi nhớ

## 3.3.13. Học thẻ ghi nhớ

|                                                                                     |                                                  |
|-------------------------------------------------------------------------------------|--------------------------------------------------|
| <b>Mã Usecase:</b> UC-13                                                            |                                                  |
| <b>Tên Usecase:</b> Học thẻ ghi nhớ                                                 |                                                  |
| <b>Mô tả:</b> Học thẻ ghi nhớ                                                       |                                                  |
| <b>Tác nhân:</b> Người dùng                                                         |                                                  |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống, truy cập vào giao diện học bộ thẻ |                                                  |
| <b>Hậu điều kiện:</b> Hệ thống cập nhật lịch sử học thẻ và lên lịch học thẻ.        |                                                  |
| <b>Luồng sự kiện chính</b>                                                          |                                                  |
| <b>Tác nhân</b>                                                                     | <b>Hệ thống</b>                                  |
| 1.Nhấn vào nút “Hiện đáp án”                                                        |                                                  |
| 2.Nhấn vào một trong các nút trạng thái ghi nhớ.                                    |                                                  |
|                                                                                     | 3. Hệ thống cập nhật lịch sử và lên lịch học thẻ |
| <b>Luồng sự kiện thay thế</b>                                                       |                                                  |
|                                                                                     |                                                  |

Bảng 14 Đặc tả usecase học thẻ ghi nhớ

## 3.3.14. Thống kê học thẻ

|                                                                            |                                                                                                                         |
|----------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|
| <b>Mã Usecase:</b> UC-14                                                   |                                                                                                                         |
| <b>Tên Usecase:</b> Thống kê học thẻ                                       |                                                                                                                         |
| <b>Mô tả:</b> Xem thống kê quá trình học tập và trạng thái các thẻ ghi nhớ |                                                                                                                         |
| <b>Tác nhân:</b> Người dùng                                                |                                                                                                                         |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống, truy cập vào bộ thẻ      |                                                                                                                         |
| <b>Hậu điều kiện:</b> Hệ thống hiển thị thông tin thống kê cho người dùng  |                                                                                                                         |
| <b>Luồng sự kiện chính</b>                                                 |                                                                                                                         |
| <b>Tác nhân</b>                                                            | <b>Hệ thống</b>                                                                                                         |
| 1.Nhấn vào nút “Stats”                                                     |                                                                                                                         |
|                                                                            | 2.Hệ thống hiển thị biểu đồ heatmap thẻ hiện quá trình học tập và biểu đồ hình tròn thể hiện trạng thái của thẻ ghi nhớ |
| <b>Luồng sự kiện thay thế</b>                                              |                                                                                                                         |

Bảng 15 Đặc tả usecase thống kê học thể

**3.3.15. Chỉnh sửa thông tin cá nhân**

|                                                                                  |                                                     |
|----------------------------------------------------------------------------------|-----------------------------------------------------|
| <b>Mã Usecase:</b> UC-15                                                         |                                                     |
| <b>Tên Usecase:</b> Chỉnh sửa thông tin cá nhân                                  |                                                     |
| <b>Mô tả:</b> Chỉnh sửa các thông tin cá nhân của người dùng.                    |                                                     |
| <b>Tác nhân:</b> Người dùng                                                      |                                                     |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống, truy cập vào giao diện cài đặt |                                                     |
| <b>Hậu điều kiện:</b> Hệ thống cập nhật thông tin của tài khoản                  |                                                     |
| <b>Luồng sự kiện chính</b>                                                       |                                                     |
| <b>Tác nhân</b>                                                                  | <b>Hệ thống</b>                                     |
| 1.Nhấn vào nút “Chỉnh sửa”                                                       |                                                     |
| 2.Nhập Họ tên                                                                    |                                                     |
| 3.Nhập Ngày sinh                                                                 |                                                     |
| 4.Chọn Giới tính                                                                 |                                                     |
| 5.Chọn Trình độ                                                                  |                                                     |
| 6.Chọn Nhu cầu                                                                   |                                                     |
| 7.Nhấn nút “Cập nhật”                                                            |                                                     |
|                                                                                  | 7.Hệ thống thông báo “Cập nhật thông tin thành công |
| <b>Luồng sự kiện thay thế</b>                                                    |                                                     |
|                                                                                  |                                                     |

Bảng 16 Đặc tả usecase chỉnh sửa thông tin cá nhân

**3.3.16. Chỉnh sửa thông tin hiển thị**

|                                                                                  |                                                             |
|----------------------------------------------------------------------------------|-------------------------------------------------------------|
| <b>Mã Usecase:</b> UC-16                                                         |                                                             |
| <b>Tên Usecase:</b> Chỉnh sửa thông tin hiển thị                                 |                                                             |
| <b>Mô tả:</b> Chỉnh sửa username và ảnh đại diện của người dùng                  |                                                             |
| <b>Tác nhân:</b> Người dùng                                                      |                                                             |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống, truy cập vào giao diện cài đặt |                                                             |
| <b>Hậu điều kiện:</b> Hệ thống cập nhật thông tin của tài khoản                  |                                                             |
| <b>Luồng sự kiện chính</b>                                                       |                                                             |
| <b>Tác nhân</b>                                                                  | <b>Hệ thống</b>                                             |
| 1.Nhấn vào icon chỉnh sửa username                                               |                                                             |
| 2.Nhập username                                                                  |                                                             |
| 3.Nhấn icon dấu tick                                                             |                                                             |
|                                                                                  | 4.Hệ thống cập nhật và hiển thị username đã cập nhật.       |
| 5.Nhấn vào icon upload ảnh đại diện                                              |                                                             |
| 6.Nhấn vào icon upload ảnh                                                       |                                                             |
| 7.Chọn ảnh từ thiết bị sử dụng.                                                  |                                                             |
| 8.Nhấn vào nút “Lưu thay đổi”                                                    |                                                             |
|                                                                                  | 9.Hệ thống cập nhật ảnh đại diện và hiển thị cho người dùng |

| Luồng sự kiện thay thế           |                                               |
|----------------------------------|-----------------------------------------------|
| 8.1.Nhấn vào nút “Chọn ảnh khác” | 9.1. Hệ thống cập nhật ảnh đại diện thất bại. |

Bảng 17 Đặc tả usecase chỉnh sửa thông tin hiển thị

**3.3.17. Thay đổi mật khẩu**

| <b>Mã Usecase:</b> UC-17                                                         |                                                       |
|----------------------------------------------------------------------------------|-------------------------------------------------------|
| <b>Tên Usecase:</b> Thay đổi mật khẩu                                            |                                                       |
| <b>Mô tả:</b> Thay đổi mật khẩu tài khoản                                        |                                                       |
| <b>Tác nhân:</b> Người dùng, quản lý                                             |                                                       |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống, truy cập vào giao diện cài đặt |                                                       |
| <b>Hậu điều kiện:</b> Hệ thống cập nhật thông tin của tài khoản                  |                                                       |
| Luồng sự kiện chính                                                              |                                                       |
| Tác nhân                                                                         | Hệ thống                                              |
| 1.Nhấn nút “Bảo mật”                                                             |                                                       |
| 2.Nhập Mật khẩu hiện tại                                                         |                                                       |
| 3.Nhập Mật khẩu mới                                                              |                                                       |
| 4.Nhập Xác nhận mật khẩu mới                                                     |                                                       |
| 5.Nhấn “Đổi mật khẩu”                                                            |                                                       |
|                                                                                  | 6.Hệ thống thông báo “Đổi mật khẩu thành công”        |
| Luồng sự kiện thay thế                                                           |                                                       |
|                                                                                  | 6.1.Hệ thống thông báo”Đổi mật khẩu không thành công” |

Bảng 18 Đặc tả usecase thay đổi mật khẩu

**3.3.18. Thêm từ vựng**

| <b>Mã Usecase:</b> UC-18                                                                                     |          |
|--------------------------------------------------------------------------------------------------------------|----------|
| <b>Tên Usecase:</b> Thêm từ vựng                                                                             |          |
| <b>Mô tả:</b> Thêm từ vựng                                                                                   |          |
| <b>Tác nhân:</b> Quản trị viên                                                                               |          |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống với vai trò quản lý và vào trang hiển thị danh sách từ vựng |          |
| <b>Hậu điều kiện:</b> Hệ thống thêm từ vào cơ sở dữ liệu                                                     |          |
| Luồng sự kiện chính                                                                                          |          |
| Tác nhân                                                                                                     | Hệ thống |
| 1.Nhấn vào nút “Thêm từ vựng”                                                                                |          |
| 2.Nhập từ vựng                                                                                               |          |
| 3.Nhập cách đọc                                                                                              |          |
| 4.Chọn loại từ vựng (Danh từ, Động từ,...)                                                                   |          |
| 5.Nhập ý nghĩa của từ vựng                                                                                   |          |
| 6.Nhập câu ví dụ                                                                                             |          |
| 7.Nhập hiragana cho câu ví dụ                                                                                |          |
| 8.Nhập ý nghĩa cho câu ví dụ                                                                                 |          |

|                                      |                                                                                 |
|--------------------------------------|---------------------------------------------------------------------------------|
| 9.Nhấn nút “Lưu”                     |                                                                                 |
|                                      | 10.Hệ thống thêm từ vựng thành công và hiển thị lại danh sách từ vựng cập nhật. |
| <b>Luồng sự kiện thay thế</b>        |                                                                                 |
| 3.1.Chọn thêm cách đọc               |                                                                                 |
| 5.1. Chọn thêm ý nghĩa của từ vựng   |                                                                                 |
| 6.1. Chọn thêm câu ví dụ cho từ vựng |                                                                                 |
|                                      | 10.1.Hệ thống thông báo lỗi “Thêm từ thất bại”                                  |

Bảng 19 Đặc tả usecase thêm từ vựng

**3.3.19. Cập nhật từ vựng**

|                                                                                                              |                                                                                     |
|--------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| <b>Mã Usecase:</b> UC-19                                                                                     |                                                                                     |
| <b>Tên Usecase:</b> Cập nhật từ vựng                                                                         |                                                                                     |
| <b>Mô tả:</b> Cập nhật thông tin từ vựng                                                                     |                                                                                     |
| <b>Tác nhân:</b> Quản trị viên                                                                               |                                                                                     |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống với vai trò quản lý và vào trang hiển thị danh sách từ vựng |                                                                                     |
| <b>Hậu điều kiện:</b> Hệ thống cập nhật nội dung từ vựng                                                     |                                                                                     |
| <b>Luồng sự kiện chính</b>                                                                                   |                                                                                     |
| <b>Tác nhân</b>                                                                                              | <b>Hệ thống</b>                                                                     |
| 1.Nhấn vào icon “3 chấm” ở cột hành động                                                                     |                                                                                     |
| 2.Nhấn vào nút “Chỉnh sửa”                                                                                   |                                                                                     |
| 3.Nhập từ vựng                                                                                               |                                                                                     |
| 4.Nhập cách đọc                                                                                              |                                                                                     |
| 5.Chọn loại từ vựng (Danh từ, Động từ,...)                                                                   |                                                                                     |
| 6.Nhập ý nghĩa của từ vựng                                                                                   |                                                                                     |
| 7.Nhập câu ví dụ                                                                                             |                                                                                     |
| 8.Nhập hiragana cho câu ví dụ                                                                                |                                                                                     |
| 9.Nhập ý nghĩa cho câu ví dụ                                                                                 |                                                                                     |
| 10.Nhấn nút “Lưu”                                                                                            |                                                                                     |
|                                                                                                              | 11.Hệ thống cập nhật từ vựng thành công và hiển thị lại danh sách từ vựng cập nhật. |
| <b>Luồng sự kiện thay thế</b>                                                                                |                                                                                     |
| 4.1.Chọn thêm cách đọc                                                                                       |                                                                                     |
| 6.1. Chọn thêm ý nghĩa của từ vựng                                                                           |                                                                                     |
| 7.1. Chọn thêm câu ví dụ cho từ vựng                                                                         |                                                                                     |
|                                                                                                              | 11.1.Hệ thống thông báo lỗi “Cập nhật từ thất bại”                                  |

Bảng 20 Đặc tả usecase cập nhật từ vựng

**3.3.20. Xóa từ**

|                                                                                                             |                                          |
|-------------------------------------------------------------------------------------------------------------|------------------------------------------|
| <b>Mã Usecase:</b> UC-20                                                                                    |                                          |
| <b>Tên Usecase:</b> Xóa từ                                                                                  |                                          |
| <b>Mô tả:</b> Xóa từ vựng và hán tự tương ứng                                                               |                                          |
| <b>Tác nhân:</b> Quản trị viên                                                                              |                                          |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống với vai trò quản lý và trang danh sách từ vựng hoặc hán tự |                                          |
| <b>Hậu điều kiện:</b> Hệ thống xóa từ tương ứng ra khỏi cơ sở dữ liệu                                       |                                          |
| <b>Luồng sự kiện chính</b>                                                                                  |                                          |
| <b>Tác nhân</b>                                                                                             | <b>Hệ thống</b>                          |
| 1.Chọn vào icon “3 chấm” ở cột hành động                                                                    |                                          |
| 2.Chọn nút xóa                                                                                              |                                          |
|                                                                                                             | 3.Hệ thống thông báo “Xóa từ thành công” |
| <b>Luồng sự kiện thay thế</b>                                                                               |                                          |
|                                                                                                             |                                          |

Bảng 21 Đặc tả usecase xóa từ

**3.3.21. Thêm hán tự**

|                                                                                                             |                                                                               |
|-------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| <b>Mã Usecase:</b> UC-21                                                                                    |                                                                               |
| <b>Tên Usecase:</b> Thêm hán tự                                                                             |                                                                               |
| <b>Mô tả:</b> Thêm hán tự                                                                                   |                                                                               |
| <b>Tác nhân:</b> Quản trị viên                                                                              |                                                                               |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống với vai trò quản lý và vào trang hiển thị danh sách hán tự |                                                                               |
| <b>Hậu điều kiện:</b> Hệ thống thêm hán tự vào cơ sở dữ liệu                                                |                                                                               |
| <b>Luồng sự kiện chính</b>                                                                                  |                                                                               |
| <b>Tác nhân</b>                                                                                             | <b>Hệ thống</b>                                                               |
| 1.Nhấn vào nút “Thêm Hán tự”                                                                                |                                                                               |
| 2.Nhập Từ vựng                                                                                              |                                                                               |
| 3.Chọn mục JLPT                                                                                             |                                                                               |
| 4.Nhập Số nét                                                                                               |                                                                               |
| 5.Nhập Bộ thủ                                                                                               |                                                                               |
| 6.Nhập Ý nghĩa                                                                                              |                                                                               |
| 7.Nhập Hán Việt                                                                                             |                                                                               |
| 8.Nhập Âm on                                                                                                |                                                                               |
| 9.Nhập Âm kun                                                                                               |                                                                               |
| 10.Nhấn nút “Lưu”                                                                                           |                                                                               |
|                                                                                                             | 11.Hệ thống thêm hán tự thành công và hiển thị lại danh sách hán tự cập nhật. |
| <b>Luồng sự kiện thay thế</b>                                                                               |                                                                               |
| 7.1.Chọn thêm Hán Việt                                                                                      |                                                                               |
| 8.1. Chọn thêm Âm on                                                                                        |                                                                               |
| 9.1. Chọn thêm Âm kun                                                                                       |                                                                               |

|  |                                                    |
|--|----------------------------------------------------|
|  | 11.1.Hệ thống thông báo lỗi “Thêm hán tự thất bại” |
|--|----------------------------------------------------|

Bảng 22 Đặc tả usecase thêm hán tự

**3.3.22. Cập nhật hán tự**

|                                                                                                             |                                                                               |
|-------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| <b>Mã Usecase:</b> UC-22                                                                                    |                                                                               |
| <b>Tên Usecase:</b> Cập nhật hán tự                                                                         |                                                                               |
| <b>Mô tả:</b> Cập nhật thông tin hán tự                                                                     |                                                                               |
| <b>Tác nhân:</b> Quản trị viên                                                                              |                                                                               |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống với vai trò quản lý và vào trang hiển thị danh sách hán tự |                                                                               |
| <b>Hậu điều kiện:</b> Hệ thống cập nhật hán tự vào cơ sở dữ liệu                                            |                                                                               |
| <b>Luồng sự kiện chính</b>                                                                                  |                                                                               |
| <b>Tác nhân</b>                                                                                             | <b>Hệ thống</b>                                                               |
| 1. Nhấn vào icon “3 chấm” ở cột hành động                                                                   |                                                                               |
| 2. Nhấn vào nút “Chỉnh sửa”                                                                                 |                                                                               |
| 3.Nhập Từ vựng                                                                                              |                                                                               |
| 4.Chọn mục JLPT                                                                                             |                                                                               |
| 5.Nhập Số nét                                                                                               |                                                                               |
| 6.Nhập Bộ thủ                                                                                               |                                                                               |
| 7.Nhập Ý nghĩa                                                                                              |                                                                               |
| 8.Nhập Hán Việt                                                                                             |                                                                               |
| 9.Nhập Âm on                                                                                                |                                                                               |
| 10.Nhập Âm kun                                                                                              |                                                                               |
| 11.Nhấn nút “Lưu”                                                                                           |                                                                               |
|                                                                                                             | 12.Hệ thống thêm hán tự thành công và hiển thị lại danh sách hán tự cập nhật. |
| <b>Luồng sự kiện thay thế</b>                                                                               |                                                                               |
| 8.1.Chọn thêm Hán Việt                                                                                      |                                                                               |
| 9.1. Chọn thêm Âm on                                                                                        |                                                                               |
| 10.1. Chọn thêm Âm kun                                                                                      |                                                                               |
|                                                                                                             | 12.1.Hệ thống thông báo lỗi “Thêm hán tự thất bại”                            |

Bảng 23 Đặc tả usecase cập nhật hán tự

**3.3.23. Xem danh sách từ**

|                                                                                   |
|-----------------------------------------------------------------------------------|
| <b>Mã Usecase:</b> UC-23                                                          |
| <b>Tên Usecase:</b> Xem danh sách từ                                              |
| <b>Mô tả:</b> Xoá danh sách từ vựng hoặc hán tự                                   |
| <b>Tác nhân:</b> Quản trị viên                                                    |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống với vai trò quản lý              |
| <b>Hậu điều kiện:</b> Hệ thống hiển thị danh sách từ vựng, hán tự theo phân trang |

| Luồng sự kiện chính      |                                                   |
|--------------------------|---------------------------------------------------|
| Tác nhân                 | Hệ thống                                          |
| 1.Chọn mục “Từ điển”     |                                                   |
| 2.Chọn danh mục Từ vựng  |                                                   |
|                          | 3.Hệ thống hiển thị danh sách từ vựng phân trang  |
| Luồng sự kiện thay thế   |                                                   |
| 2.1 Chọn danh mục Hán tự |                                                   |
|                          | 3.1 Hệ thống hiển thị danh sách hán tự phân trang |

Bảng 24 Đặc tả usecase xem danh sách từ

## 3.3.24. Xem danh sách người dùng

| <b>Mã Usecase:</b> UC-24                                                |                                                     |
|-------------------------------------------------------------------------|-----------------------------------------------------|
| <b>Tên Usecase:</b> Xem danh sách người dùng                            |                                                     |
| <b>Mô tả:</b> Xem danh sách người dùng                                  |                                                     |
| <b>Tác nhân:</b> Quản trị viên                                          |                                                     |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống với vai trò quản lý    |                                                     |
| <b>Hậu điều kiện:</b> Hệ thống hiển thị danh sách người dùng phân trang |                                                     |
| Luồng sự kiện chính                                                     |                                                     |
| Tác nhân                                                                | Hệ thống                                            |
| 1.Chọn mục “Người dùng”                                                 |                                                     |
|                                                                         | 2.Hệ thống hiển thị danh sách người dùng phân trang |
| Luồng sự kiện thay thế                                                  |                                                     |
|                                                                         |                                                     |

Bảng 25 Đặc tả usecase xem danh sách từ

## 3.3.25. Cập nhật thông tin tài khoản Quản trị viên

| <b>Mã Usecase:</b> UC-25                                                                         |                                         |
|--------------------------------------------------------------------------------------------------|-----------------------------------------|
| <b>Tên Usecase:</b> Cập nhật thông tin tài khoản quản trị viên                                   |                                         |
| <b>Mô tả:</b> Cập nhật các thông tin tài khoản quản trị viên như avatar, thông tin quản trị viên |                                         |
| <b>Tác nhân:</b> Quản trị viên                                                                   |                                         |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống với vai trò quản lý                             |                                         |
| <b>Hậu điều kiện:</b> Hệ thống cập nhật thông tin tài khoản quản trị viên                        |                                         |
| Luồng sự kiện chính                                                                              |                                         |
| Tác nhân                                                                                         | Hệ thống                                |
| 1.Chọn mục “Cài đặt”                                                                             |                                         |
| 2. Nhập Họ và tên                                                                                |                                         |
| 3. Nhập Username                                                                                 |                                         |
| 4. Nhấn vào icon camera                                                                          |                                         |
| 5. Chọn ảnh avatar từ thiết bị                                                                   |                                         |
| 6. Nhấn “Lưu thay đổi”                                                                           |                                         |
|                                                                                                  | 7.Hệ thống cập nhật thông tin tài khoản |

|                               |                                                                  |
|-------------------------------|------------------------------------------------------------------|
|                               | thành công và hiển thị giao diện thông tin tài khoản đã cập nhật |
| <b>Luồng sự kiện thay thế</b> |                                                                  |
|                               |                                                                  |

Bảng 26 Đặc tả usecase cập nhật thông tin tài khoản Quản trị viên

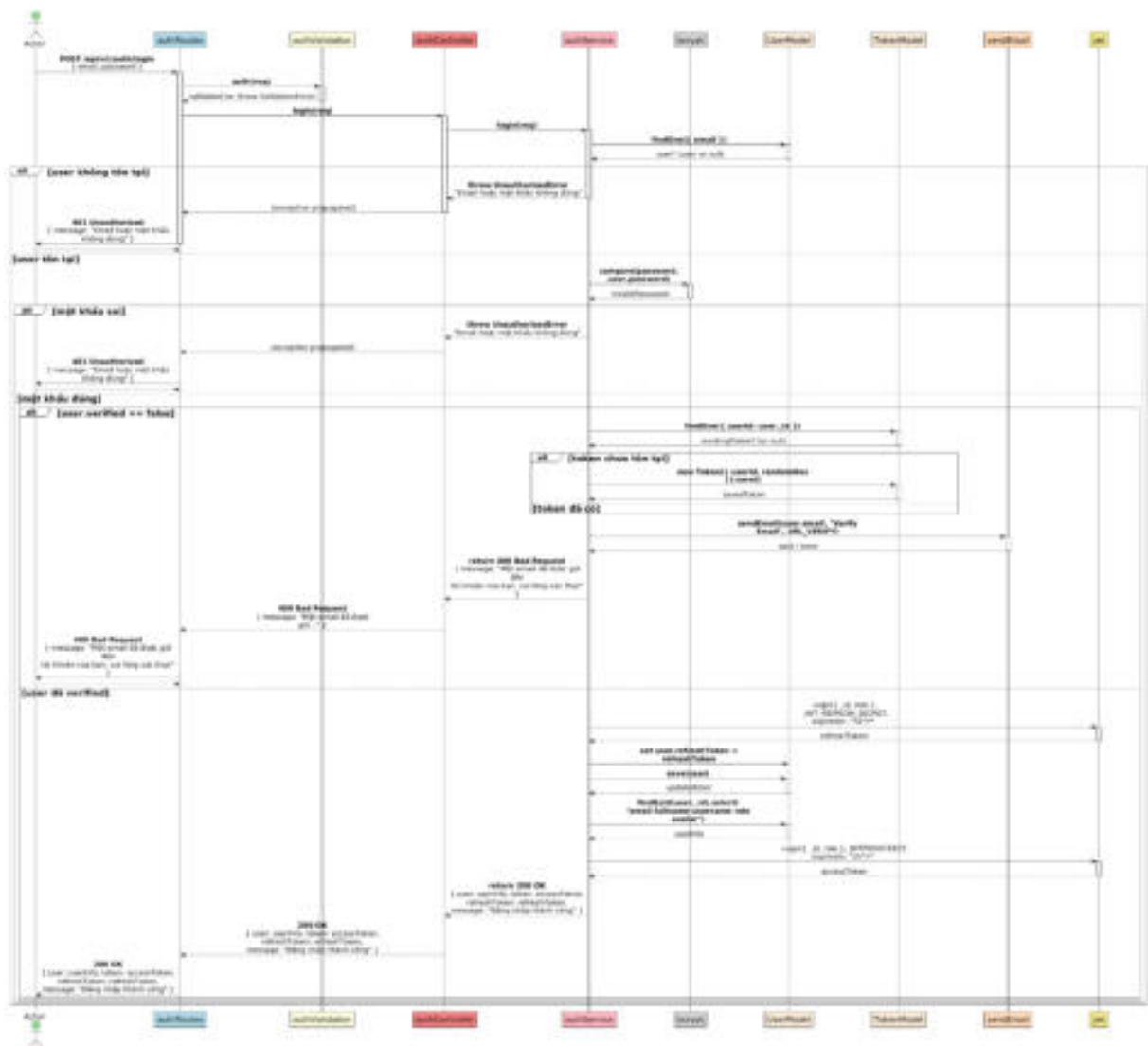
### 3.3.26. Xem thống kê

|                                                                      |                                           |
|----------------------------------------------------------------------|-------------------------------------------|
| <b>Mã Usecase:</b> UC-26                                             |                                           |
| <b>Tên Usecase:</b> Thống kê                                         |                                           |
| <b>Mô tả:</b> Thống kê từ vựng, hán tự và người dùng của ứng dụng    |                                           |
| <b>Tác nhân:</b> Quản trị viên                                       |                                           |
| <b>Tiền điều kiện:</b> Đã đăng nhập vào hệ thống với vai trò quản lý |                                           |
| <b>Hậu điều kiện:</b> Hiển thị các biểu đồ tương ứng                 |                                           |
| <b>Luồng sự kiện chính</b>                                           |                                           |
| <b>Tác nhân</b>                                                      | <b>Hệ thống</b>                           |
| 1.Chọn vào mục “Dashboard”                                           |                                           |
|                                                                      | 2.Hệ thống hiển thị các biểu đồ tương ứng |
| <b>Luồng sự kiện thay thế</b>                                        |                                           |
|                                                                      |                                           |

Bảng 27 Đặc tả usecase xem thống kê

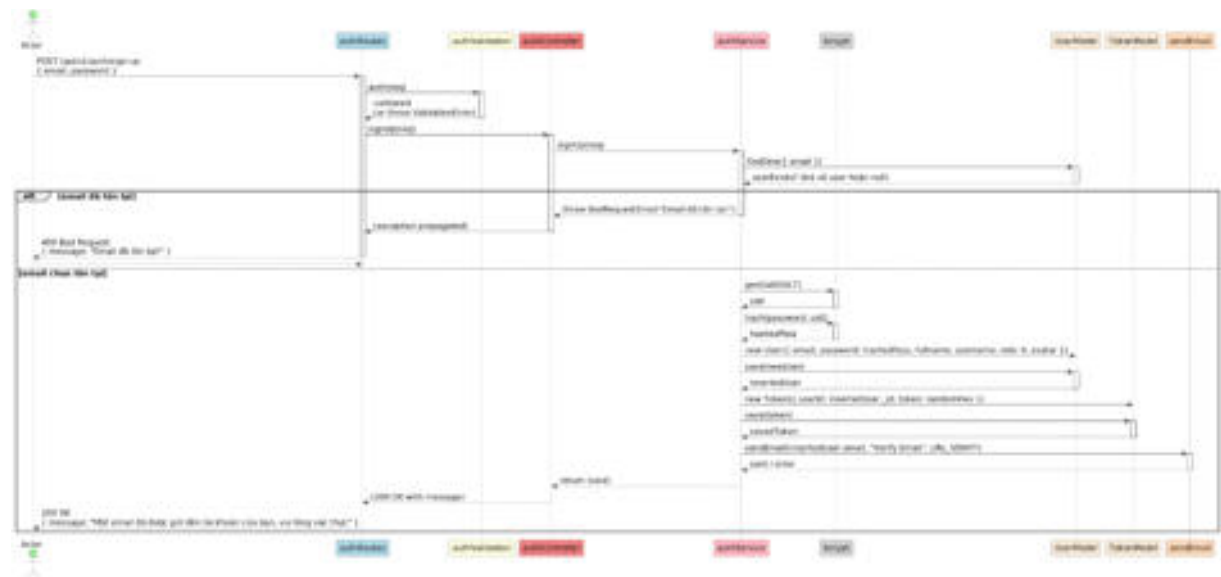
## 3.4. Biểu đồ tuần tự

### 3.4.1. Chức năng đăng nhập



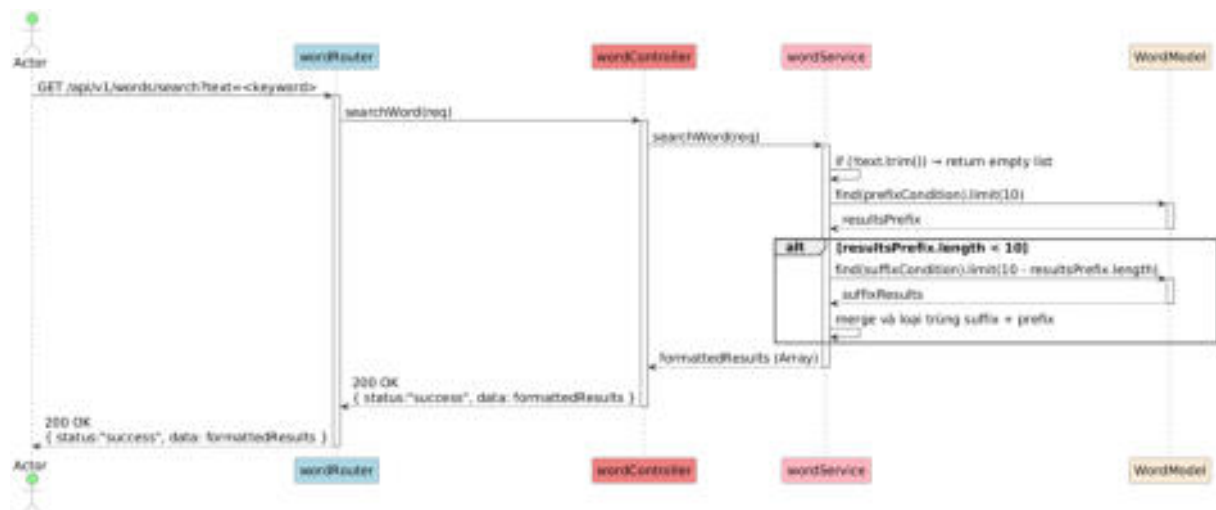
Hình 15 Biểu đồ tuần tự chức năng đăng nhập

### 3.4.2. Chức năng đăng ký



Hình 16 Biểu đồ tuần tự chức năng đăng ký

### 3.4.3. Chức năng tra cứu từ vựng



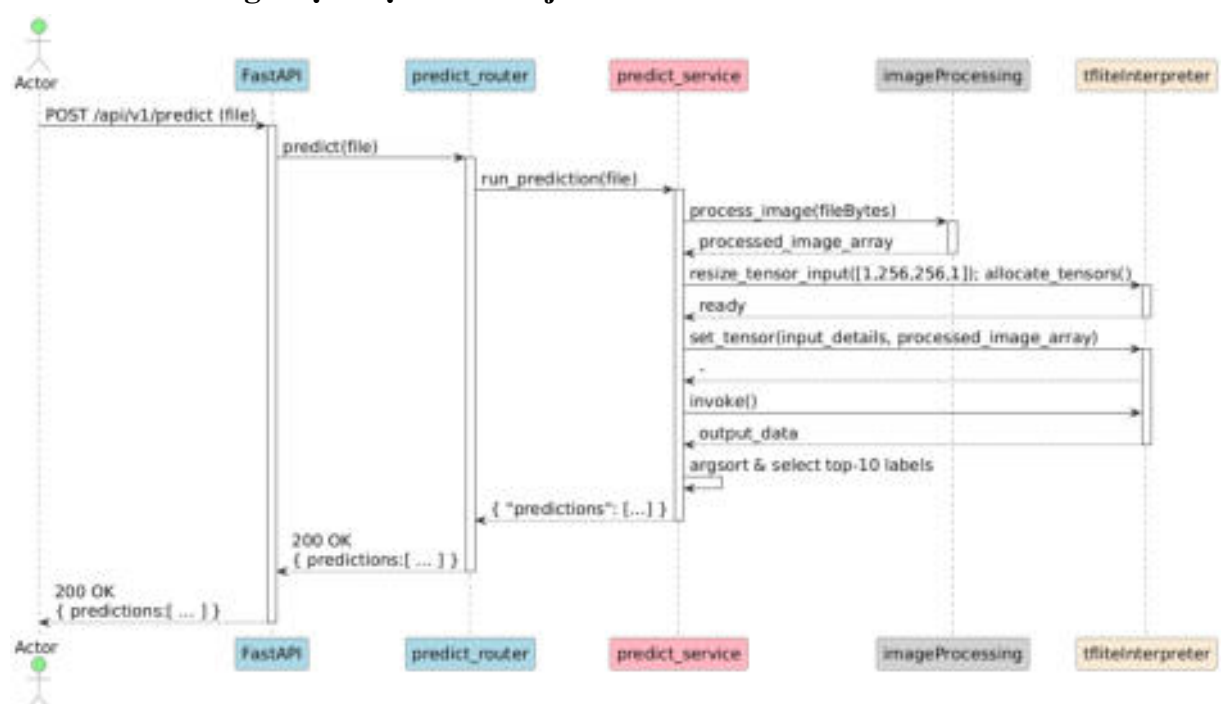
Hình 17 Biểu đồ tuần tự chức năng tra cứu từ vựng

### 3.4.4. Chức năng OCR hình ảnh



Hình 18 Biểu đồ tuần tự chức năng OCR hình ảnh

### 3.4.5. Chức năng nhận diện chữ kanji



Hình 19 Biểu đồ tuần tự chức năng nhận diện chữ kanji



### 3.4.8. Chức năng học thẻ ghi nhớ



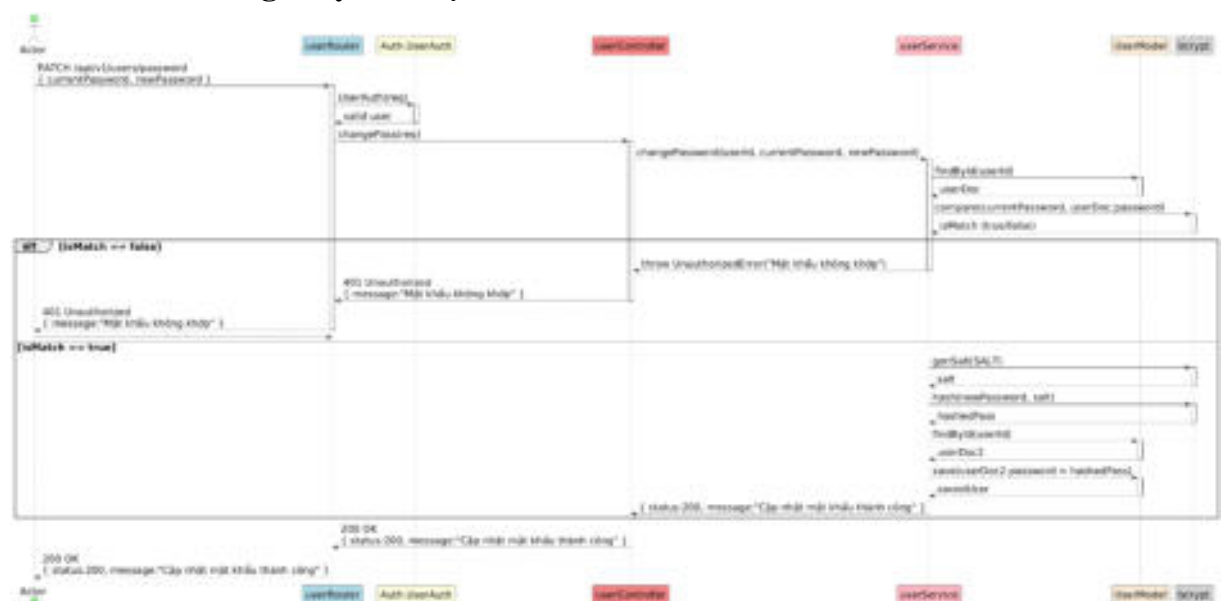
Hình 22 Biểu đồ tuần tự chức năng học thẻ ghi nhớ

### 3.4.9. Chức năng cập nhật thông tin người dùng



Hình 23 Biểu đồ tuần tự chức năng cập nhật thông tin người dùng

### 3.4.10. Chức năng thay đổi mật khẩu



Hình 24 Biểu đồ tuần tự chức năng thay đổi mật khẩu

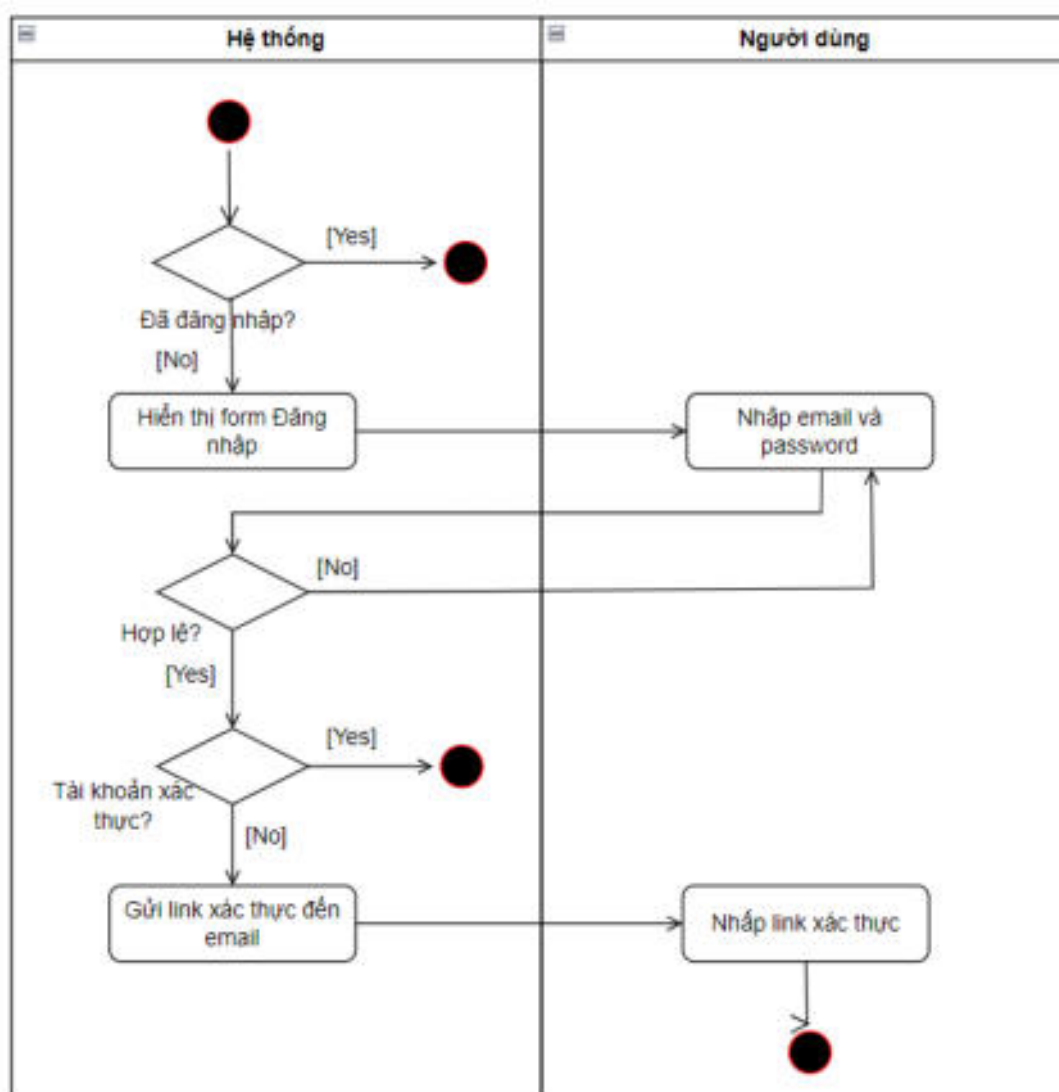
### 3.4.11. Chức năng thống kê



Hình 25 Biểu đồ tuần tự chức năng thống kê

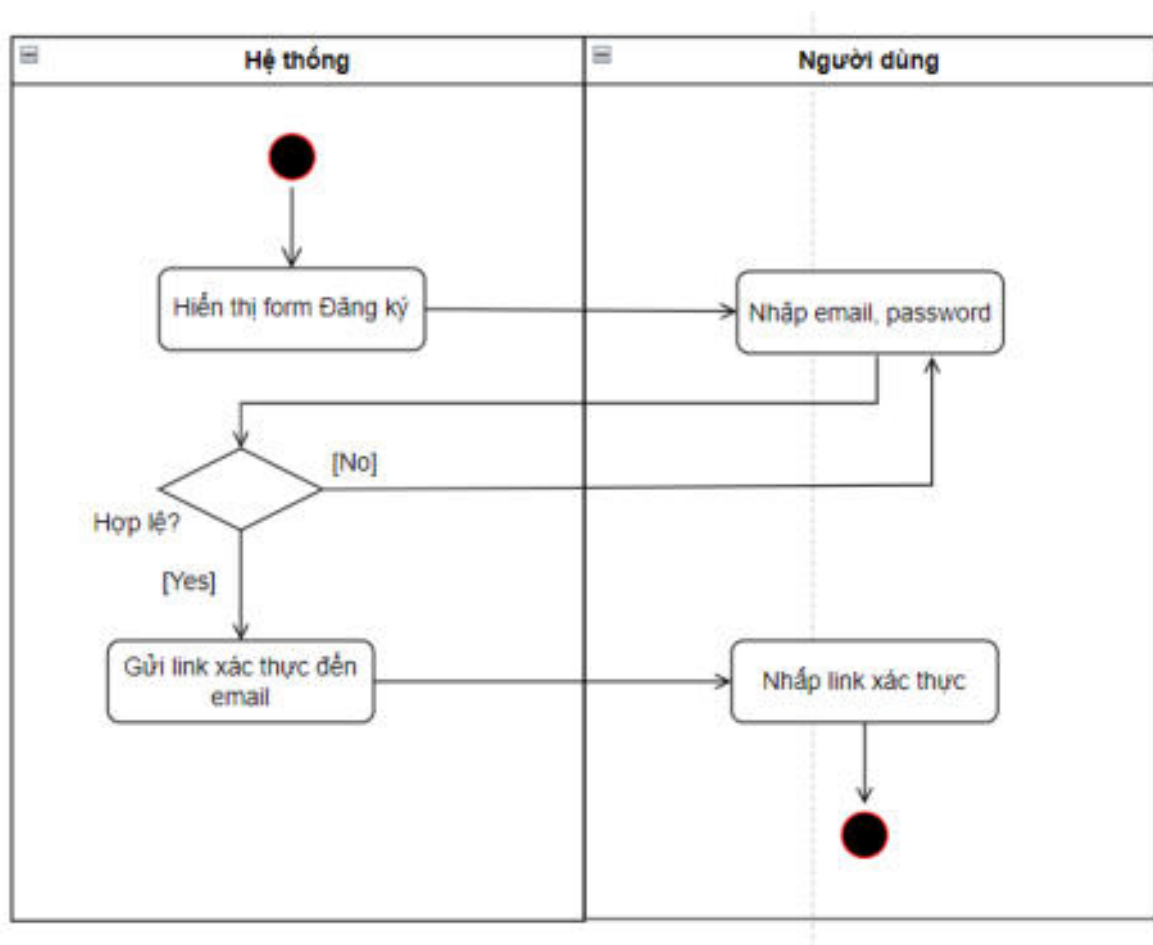
## 3.5. Biểu đồ hoạt động

### 3.5.1. Chức năng đăng nhập



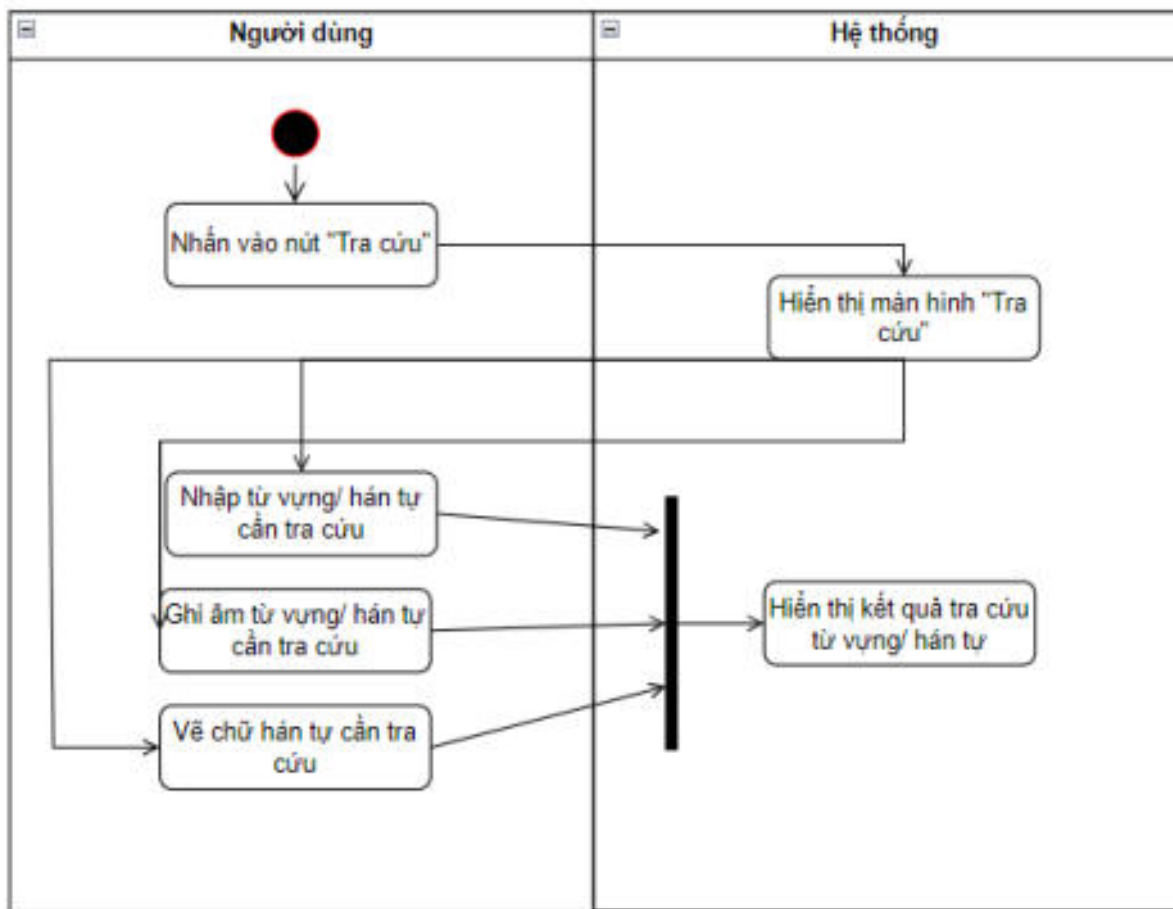
Hình 26 Sơ đồ hoạt động chức năng đăng nhập

### 3.5.2. Chức năng đăng ký



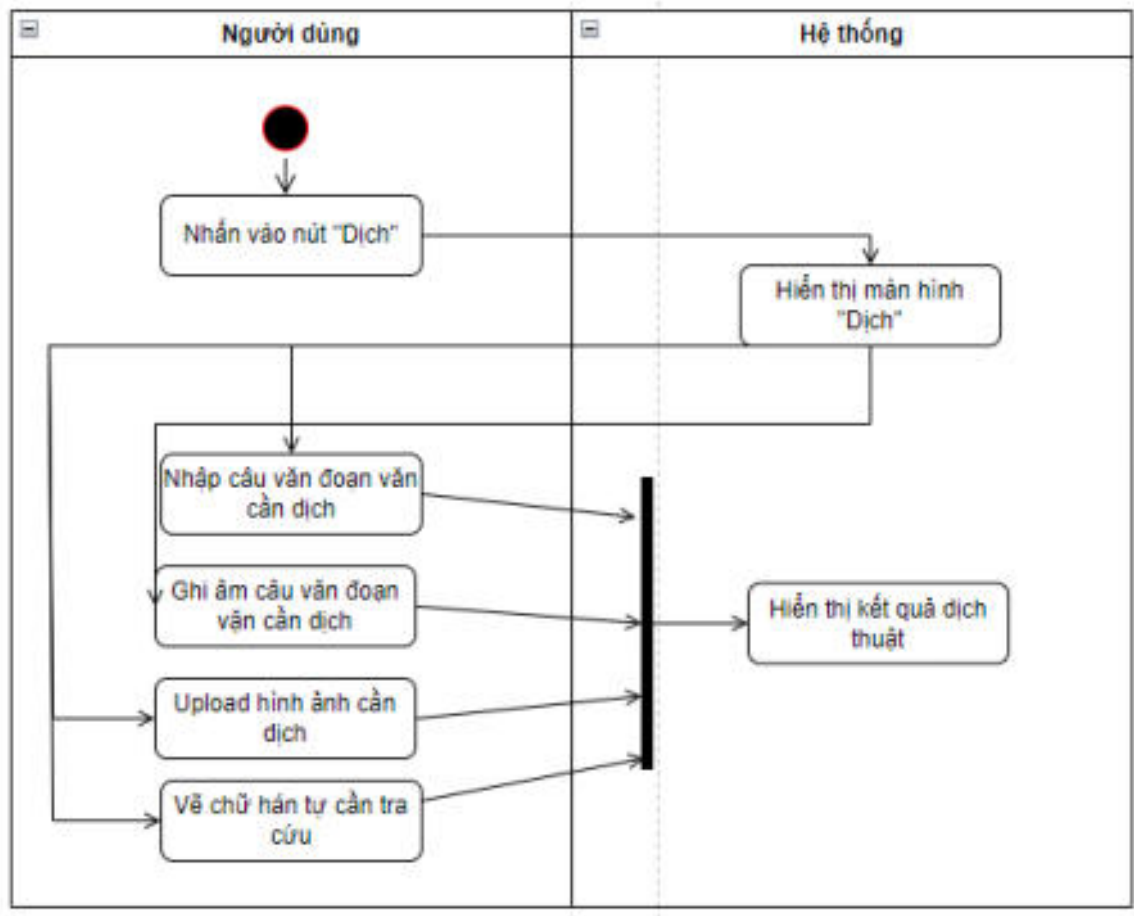
Hình 27 Sơ đồ hoạt động chức năng đăng ký

### 3.5.3. Chức năng tra cứu từ vựng



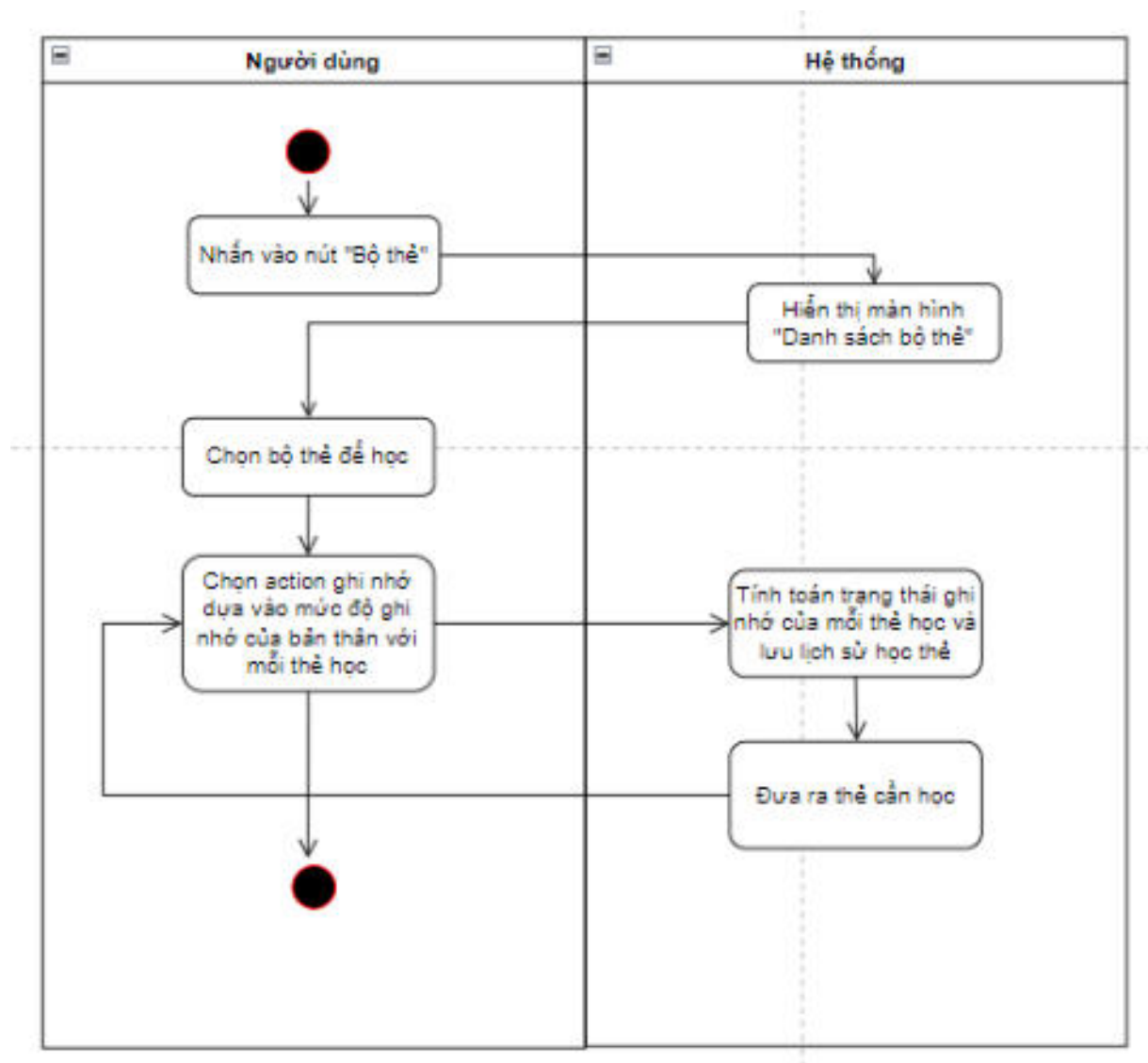
Hình 28 Sơ đồ hoạt động chức năng tra cứu từ vựng

### 3.5.4. Chức năng dịch câu



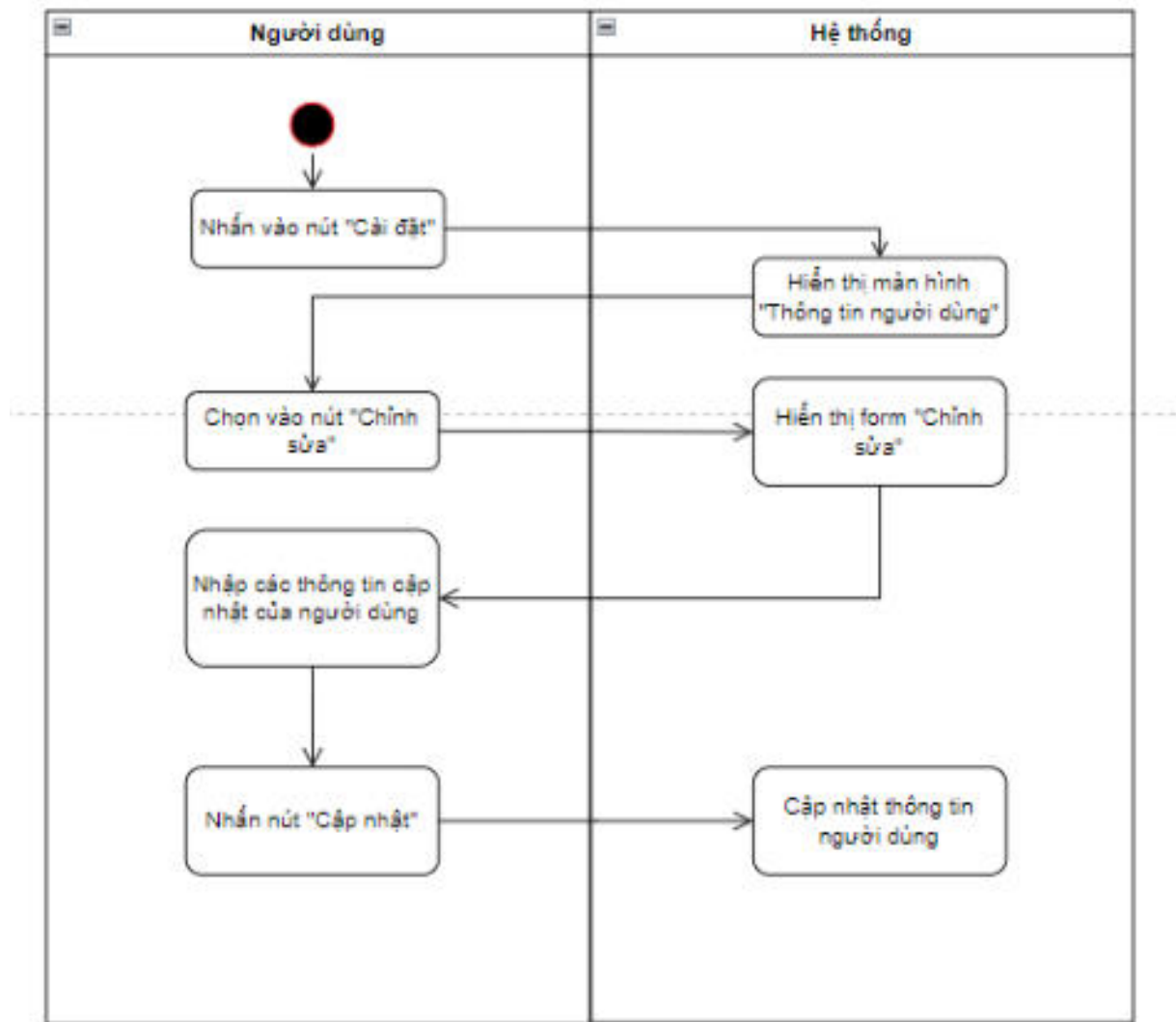
Hình 29 Sơ đồ hoạt động chức năng Dịch câu

### 3.5.5. Chức năng học thẻ ghi nhớ



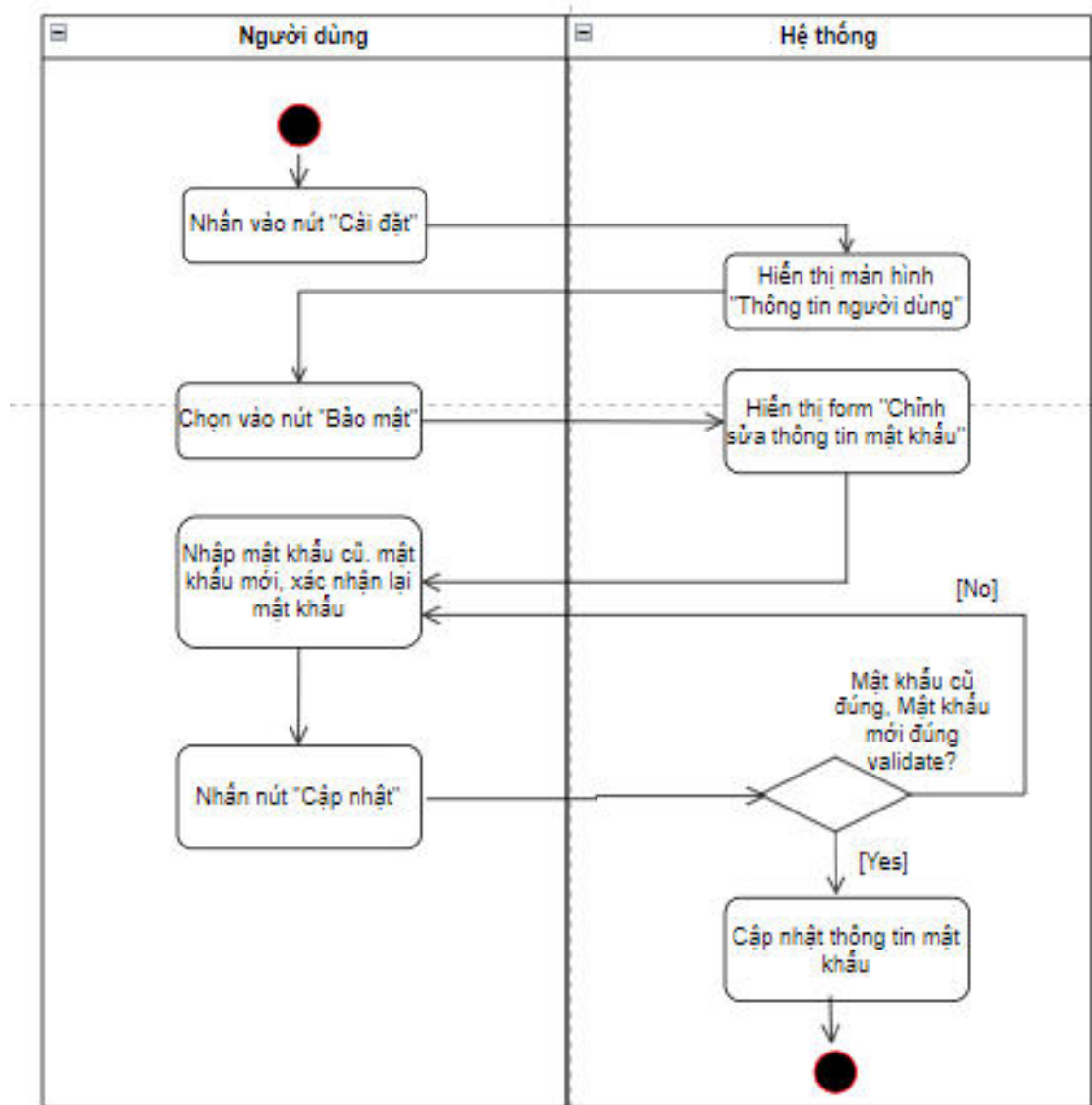
Hình 30 Sơ đồ hoạt động chức năng Học thẻ ghi nhớ

### 3.5.6. Chức năng cập nhật thông tin người dùng



Hình 31 Sơ đồ hoạt động chức năng Cập nhật thông tin người dùng

### 3.5.7. Chức năng thay đổi mật khẩu



Hình 32 Sơ đồ hoạt động chức năng Thay đổi mật khẩu

### 3.6. Cơ sở dữ liệu



Hình 33 Cơ sở dữ liệu

#### 3.6.1. Phân tích bảng cơ sở dữ liệu

##### 3.6.1.1. User

| Tên trường | Kiểu dữ liệu | Mô tả                                |
|------------|--------------|--------------------------------------|
| id         | ObjectId     | Mã người dùng                        |
| email      | Varchar      | Địa chỉ email                        |
| password   | Varchar      | Mật khẩu đã hash                     |
| fullname   | Varchar      | Tên đầy đủ                           |
| username   | Varchar      | Tên hiển thị                         |
| role       | Integer      | Vai trò: admin (1), user(0)          |
| birthday   | Timestamp    | Ngày tháng năm sinh                  |
| sex        | Boolean      | Giới tính: Nam (1), Nữ (0)           |
| level      | Integer      | Trình độ tiếng Nhật: N1(1)           |
| demand     | Varchar      | Nhu cầu người dùng                   |
| verified   | Boolean      | Người dùng đã được xác thực hay chưa |

|              |           |                                            |
|--------------|-----------|--------------------------------------------|
| avatar       | Varchar   | Đường dẫn tới ảnh đại diện                 |
| created_at   | Timestamp | Thời gian tài khoản được tạo               |
| updated_at   | Timestamp | Thời gian gần nhất tài khoản được cập nhật |
| refreshToken | Varchar   | Refresh token                              |

Bảng 28 Phân tích bảng User

### 3.6.1.2. Words

| Tên trường | Kiểu dữ liệu | Mô tả                          |
|------------|--------------|--------------------------------|
| id         | ObjectId     | Mã từ                          |
| text       | Varchar      | Từ                             |
| hiragana   | Varchar[]    | Hiragana của từ                |
| meaning    | Json[]       | Danh sách nghĩa của từ         |
| examples   | Json[]       | Danh sách ví dụ của từ         |
| kanji      | Json[]       | Danh sách từ kanji có trong từ |
| romanji    | Varchar[]    | Romanji của từ                 |

Bảng 29 Phân tích bảng Words

### 3.6.1.3. Kanjis

| Tên trường   | Kiểu dữ liệu | Mô tả                 |
|--------------|--------------|-----------------------|
| id           | ObjectId     | Mã hán tự             |
| text         | Varchar      | Hán tự                |
| phonetic     | Varchar[]    | Danh sách âm Hán Việt |
| onyomi       | Varchar[]    | Danh sách âm On       |
| kunyomi      | Varchar[]    | Danh sách âm Kun      |
| strokes      | Integer      | Sô nét vẽ             |
| jlpt_level   | Integer      | Cấp độ JLPT           |
| meaning      | Varchar      | Ý nghĩa               |
| romaji       | Varchar[]    | Danh sách romaji      |
| compositions | ObjectId[]   | Danh sách bộ thủ      |

Bảng 30 Phân tích bảng Kanjis

### 3.6.1.4. Compositions

| Tên trường | Kiểu dữ liệu | Mô tả       |
|------------|--------------|-------------|
| id         | ObjectId     | Mã bộ       |
| raw_text   | Varchar      | Bộ chữ      |
| phonetic   | Varchar      | Âm Hán Việt |

Bảng 31 Phân tích bảng Compositions

## 3.6.1.5. Decks

| Tên trường   | Kiểu dữ liệu | Mô tả                         |
|--------------|--------------|-------------------------------|
| id           | ObjectId     | Mã bộ thẻ                     |
| name         | Varchar      | Tên bộ thẻ                    |
| cardQuantity | Integer      | Số lượng thẻ mới trong 1 ngày |
| createdBy    | ObjectId     | Mã người tạo ra bộ thẻ        |
| createdAt    | Timestamp    | Ngày tạo bộ thẻ               |
| updatedAt    | Timestamp    | Ngày cập nhật bộ thẻ          |

Bảng 32 Phân tích bảng Decks

## 3.6.1.6. Cards

| Tên trường     | Kiểu dữ liệu | Mô tả                                                        |
|----------------|--------------|--------------------------------------------------------------|
| id             | ObjectId     | Mã thẻ ghi nhớ                                               |
| deckId         | ObjectId     | Mã bộ thẻ mà thẻ này thuộc về                                |
| word           | Varchar      | Từ vựng                                                      |
| sentence       | Varchar      | Câu ví dụ                                                    |
| reading        | Varchar      | Cách đọc                                                     |
| meaning        | Varchar      | Ý nghĩa                                                      |
| difficulty     | Float        | Phản ánh độ khó tiềm ẩn của nội dung thẻ                     |
| due            | Timestamp    | Ngày thẻ này sẽ đến hạn ôn tập tiếp theo                     |
| elapsed_days   | Integer      | Số ngày kể từ lần cuối thẻ được ôn tập                       |
| lapses         | Integer      | Số lần thẻ bị quên hoặc nhớ sai                              |
| last_review    | Timestamp    | Ngày ôn tập gần nhất, nếu có                                 |
| reps           | Integer      | Tổng số lần thẻ đã được ôn tập                               |
| scheduled_days | Integer      | Khoảng thời gian dự kiến cho lần ôn tập tiếp theo            |
| stability      | Float        | Thước đo mức độ thông tin đã được ghi nhớ                    |
| state          | Integer      | Trạng thái hiện tại của thẻ (Mới, Đang học, Ôn tập, Học lại) |
| createdAt      | Timestamp    | Ngày tạo thẻ                                                 |
| updatedAt      | Timestamp    | Ngày cập nhật thẻ                                            |

Bảng 33 Phân tích bảng Cards

**3.6.1.7. ReviewLogs**

| Tên trường        | Kiểu dữ liệu | Mô tả                                                                                       |
|-------------------|--------------|---------------------------------------------------------------------------------------------|
| id                | ObjectId     | Mã log học tập                                                                              |
| cardId            | ObjectId     | Mã thẻ ghi nhớ                                                                              |
| difficulty        | Float        | Độ khó của thẻ trước khi ôn tập                                                             |
| due               | Timestamp    | Ngày của lần lên lịch ôn tập gần nhất                                                       |
| elapsed_days      | Integer      | Số ngày đã trôi qua kể từ lần ôn tập trước                                                  |
| last_elapsed_days | Integer      | Số ngày giữa hai lần ôn tập gần nhất                                                        |
| rating            | Integer      | Đánh giá của lần ôn tập, có thể là "Lại" (Again), "Khó" (Hard), "Tốt" (Good), "Dễ" (Easy)   |
| review            | Timestamp    | Ngày thực hiện lần ôn tập này                                                               |
| scheduled_days    | Integer      | Số ngày cho đến lần ôn tập tiếp theo theo kế hoạch                                          |
| stability         | Float        | Mức độ ổn định của thẻ trước khi thực hiện ôn tập                                           |
| state             | Integer      | Trạng thái của thẻ trong quá trình ôn tập, có thể là "Mới", "Đang học", "Ôn tập", "Học lại" |
| createdAt         | Timestamp    | Ngày tạo log                                                                                |
| updatedAt         | Timestamp    | Ngày cập nhật log                                                                           |

Bảng 34 Phân tích bảng ReviewLogs

## CHƯƠNG 4. TRIỂN KHAI MÔ HÌNH VÀ THUẬT TOÁN

### 4.1. Phát biểu bài toán

#### 4.1.1. Dịch thuật

Dự án xây dựng từ điển Nhật-Việt nhằm cung cấp một công cụ hỗ trợ dịch thuật hai chiều giữa tiếng Nhật và tiếng Việt. Hệ thống không chỉ giúp tra cứu từ vựng mà còn cung cấp chức năng dịch câu, đoạn văn một cách chính xác và tự nhiên. Các yêu cầu cụ thể bao gồm:

- **Chức năng dịch thuật:** Hỗ trợ dịch từ tiếng Nhật sang tiếng Việt và ngược lại, đảm bảo chất lượng dịch thuật cho cả từ ngữ và ngữ cảnh.
- **Tính chính xác:** Duy trì độ chính xác cao trong dịch thuật, xử lý tốt các trường hợp như từ đa nghĩa, cấu trúc câu phức tạp, và các yếu tố văn hóa của từng ngôn ngữ.
- **Khả năng mở rộng:** Cung cấp khả năng cải tiến thông qua việc bổ sung dữ liệu, huấn luyện lại mô hình, và tích hợp các tính năng thông minh trong tương lai.

Hệ thống hướng tới việc trở thành một công cụ dịch thuật toàn diện, giúp người dùng giao tiếp và học tập hiệu quả hơn giữa hai ngôn ngữ.

Nhìn chung, Input và Output của bài toán này sẽ như sau:

- **Input:** Một câu hay một đoạn văn tiếng Nhật hoặc tiếng Việt.
- **Output:** Bản dịch tương ứng.

#### 4.1.2. Thẻ ghi nhớ

Trong dự án này, một trong những tính năng quan trọng là hỗ trợ người dùng học từ vựng hiệu quả thông qua hệ thống flashcard. Mục tiêu của bài toán là thiết kế một hệ thống giúp người dùng học từ vựng tiếng Nhật một cách dễ dàng và hiệu quả, đồng thời tối ưu hóa quá trình ôn luyện thông qua việc sử dụng thuật toán FSRS (Forgotten Spaced Repetition System).

**Bài toán đặt ra là:** Làm thế nào để xây dựng một hệ thống flashcard, trong đó người dùng có thể học và ôn luyện từ vựng Nhật-Việt, đồng thời thuật toán FSRS sẽ tối ưu hóa thời gian ôn luyện và nâng cao khả năng ghi nhớ của người dùng dựa trên mức độ quên của từng từ. Điều này đòi hỏi hệ thống phải theo dõi tiến trình học của người dùng, đánh giá mức độ ghi nhớ từ vựng, và điều chỉnh lịch ôn luyện sao cho phù hợp.

### 4.2. Triển khai mô hình và thuật toán

#### 4.2.1. Triển khai mô hình Transformer.

##### 4.2.1.1. Cài đặt model

Model Transformer được cài đặt sử dụng thư viện Tensorflow 2.18.0, huấn luyện trên Google Colab. Sau đó model này được deploy bằng FastAPI để client có thể sử

dụng qua Rest API hoặc qua kết nối socket giúp nâng cao trải nghiệm người dùng.

#### 4.2.1.2. Dữ liệu huấn luyện

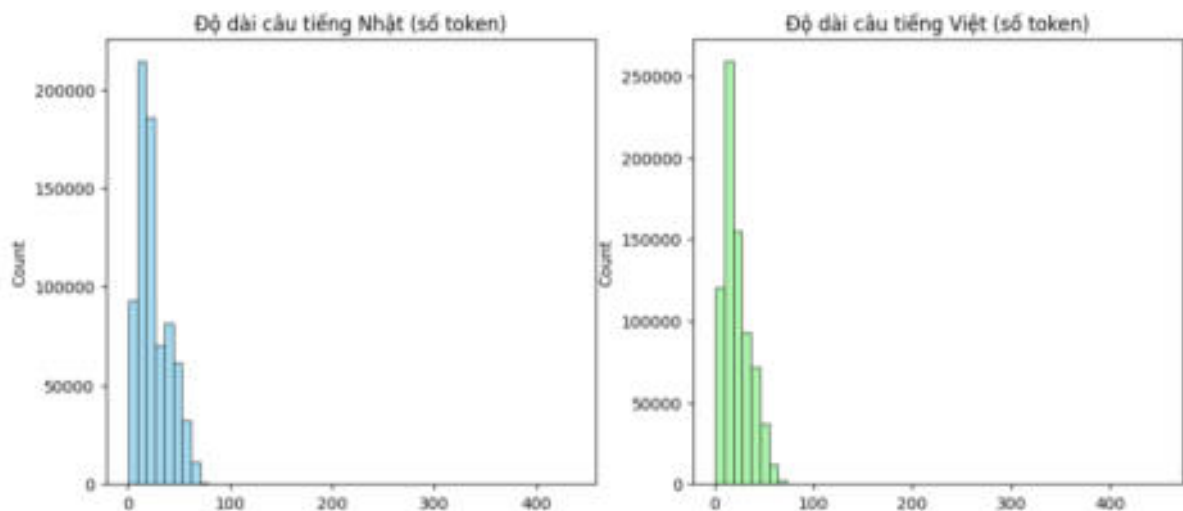
Để huấn luyện mô hình Transformer cho tác vụ dịch thuật giữa tiếng Nhật và tiếng Việt, tôi đã thu thập một bộ dữ liệu lớn với khoảng 900,000 câu tiếng Nhật từ các sách tiểu thuyết. Các câu này được chọn lựa dựa trên độ dài từ 5 đến 100 từ (tối đa 64 token), đảm bảo tính đa dạng và phong phú về cấu trúc câu và từ vựng, giúp mô hình học được nhiều mẫu câu khác nhau.

Sau khi thu thập dữ liệu, tôi tiến hành tiền xử lý bộ dữ liệu bằng cách loại bỏ các ký tự đặc biệt không cần thiết, như dấu câu đặc biệt, ký tự lạ, hoặc các yếu tố có thể gây nhiễu trong quá trình huấn luyện mô hình. Quá trình này giúp giảm thiểu sự phức tạp của bộ dữ liệu, đồng thời tăng tính hiệu quả của mô hình trong việc học và dịch.

Tiếp theo, tôi sử dụng Google Dịch để dịch các câu tiếng Nhật sang tiếng Việt. Dù việc sử dụng Google Dịch có thể tạo ra một số sai lệch nhỏ trong bản dịch, nhưng bộ dữ liệu dịch tự động này vẫn đảm bảo tính chính xác và sự liên kết giữa hai ngôn ngữ, từ đó cung cấp một cơ sở vững chắc cho việc huấn luyện mô hình dịch tự động. Hơn nữa khi training model, mô hình sẽ học dần dần từ các token trước vì nên sau quá trình training model sẽ có thể cho ra được kết quả training có thể có văn phong hay hơn Google dịch.

Bộ dữ liệu này đã được chuẩn bị để phục vụ quá trình huấn luyện mô hình Transformer, giúp mô hình học được các mối quan hệ ngữ nghĩa và cú pháp giữa tiếng Nhật và tiếng Việt, từ đó cung cấp kết quả dịch thuật chính xác và tự nhiên.

Tuy nhiên vì hạn chế tài nguyên training, chúng tôi đã phải cắt độ dài một dòng dữ liệu xuống còn 64 token. Điều này sẽ làm mất ngữ cảnh của câu đi đáng kể từ đó dẫn đến giảm độ chính xác của các câu trong văn bản dịch. Vì vậy điều tôi cần làm bây giờ là cắt mỗi dòng dữ liệu xuống còn 64 token nhưng không bị cắt ngang câu trên dòng dữ liệu. Điều này giúp giữ được ngữ cảnh trong 1 câu.



### Hình 34 Tổng quan dataset

Ở bảng trên ta có thể thấy số lượng câu tiếng Nhật và câu tiếng Việt có số token > 64 ở số ít. Điều này đảm bảo rằng trong quá trình training khi đặt số token tối đa được training trên mỗi dòng dữ liệu là MAX\_TOKENS=64 thì ngữ cảnh câu sẽ không bị cắt mất trong quá trình training. Điều này giúp cho model có thể học và hiểu ngữ cảnh tốt hơn.

#### 4.2.1.3. Tokenizer

Để xây dựng bộ tokenizer cho cả tiếng Nhật và tiếng Việt, tôi đã tiến hành một quy trình tỉ mỉ và tự xây dựng dựa trên bộ dữ liệu đã thu thập. Đối với tiếng Nhật, tôi bắt đầu bằng việc sử dụng công cụ phân tách từ Mecab, một công cụ nổi tiếng trong việc phân tích và tách câu tiếng Nhật thành các đơn vị từ (token). Mecab giúp tôi tách các câu tiếng Nhật thành các từ vựng và cụm từ có ý nghĩa riêng biệt. Sau đó, tôi sử dụng kết quả này để huấn luyện một BERT Tokenizer, với mục tiêu tối ưu hóa việc xử lý và mã hóa các từ vựng tiếng Nhật, giúp mô hình có thể hiểu và xử lý ngữ nghĩa của văn bản một cách chính xác.

Với tiếng Việt, tôi sử dụng công cụ Underthesea, một thư viện phân tích ngôn ngữ tự nhiên mạnh mẽ cho tiếng Việt, để phân tách văn bản thành các từ có nghĩa. Underthesea cho phép tôi chia câu tiếng Việt thành các đơn vị từ vựng phù hợp với ngữ cảnh của từng câu, giúp mô hình hiểu rõ cấu trúc ngữ pháp và ngữ nghĩa của ngôn ngữ này. Sau khi hoàn thành bước phân tách, tôi tiếp tục huấn luyện bộ BERT Tokenizer cho tiếng Việt, nhằm tạo ra một bộ tokenizer mạnh mẽ, có khả năng chuyển đổi các từ vựng tiếng Việt thành các mã số token để mô hình có thể dễ dàng xử lý.

Quá trình này giúp tôi xây dựng được hai bộ tokenizer đặc thù cho mỗi ngôn ngữ, tối ưu hóa khả năng hiểu và dịch các câu văn trong tiếng Nhật và tiếng Việt, từ đó nâng cao hiệu quả của mô hình dịch thuật Transformer.



Hình 35 Xử lý tokenizer

#### 4.2.1.4. Huấn luyện model

Quá trình huấn luyện mô hình được chia thành hai giai đoạn chính:

- **Giai đoạn 1:** Trong giai đoạn đầu tiên, mô hình được huấn luyện với optimizer Adam, một thuật toán tối ưu hóa hiệu quả và phổ biến trong các bài toán học sâu. Learning rate được thiết lập ở mức 0.0001, giúp mô hình học một cách ổn định mà không gặp phải hiện tượng gradient explosion hoặc gradient vanishing. Với learning rate này, mô hình có thể điều chỉnh các trọng số một cách chậm rãi và hiệu quả trong giai đoạn đầu của quá trình huấn luyện.
- **Giai đoạn 2:** Sau 80 epoch đầu tiên, khi mô hình đã học được một phần quan trọng của các đặc trưng trong dữ liệu, tôi giảm learning rate xuống 0.00001 để giúp mô hình tinh chỉnh các trọng số một cách chính xác hơn. Việc giảm learning rate này giúp quá trình tối ưu hóa đạt được sự ổn định cao hơn và tránh việc bỏ qua các tối ưu cục bộ quan trọng khi mô hình gần đạt được điểm tối ưu toàn cục.

Việc chia quá trình huấn luyện thành hai giai đoạn với các mức learning rate khác nhau giúp mô hình có thể học một cách hiệu quả và tối ưu hơn trong cả giai đoạn huấn luyện ban đầu lẫn giai đoạn tinh chỉnh.

#### 4.2.1.5. Kết quả và đánh giá mô hình.

Để phân tích và đánh giá khả năng dịch thuật từ tiếng Nhật sang tiếng Việt của mô

hình, tôi sẽ đánh giá dựa trên 12 cặp câu/ đoạn mẫu với độ phức tạp và chủ đề khác nhau. Việc đánh giá sẽ tập trung vào 3 khía cạnh chính:

- Độ chính xác nội dung (Adequacy): Mô hình có giữ nguyên và truyền tải đầy đủ ý nghĩa gốc không?
- Tính tự nhiên (Fluency): Kết quả dịch có trôi chảy, phù hợp ngữ văn tiếng Việt không?
- Fidelity (Trung thành phong cách): Mô hình có nắm bắt được sắc thái, mức độ trang trọng hay cảm xúc của nguyên tác hay không?

Kết quả thực thể của mô hình dịch:

| STT | Độ dài   | Chủ đề               | Nguyên bản                                                                       | Kết quả                                                                                                                                                                                                                                        | Nhận xét                                              |
|-----|----------|----------------------|----------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------|
| 1   | Rất ngắn | Thời tiết, động vật  | 今日はいい天気だ。猫が窓辺で寝ている。                                                              | Thời tiết hôm nay đẹp. Một chú mèo ngủ bên cửa sổ.                                                                                                                                                                                             | Đúng ý, trôi chảy, tự nhiên.                          |
| 2   | Ngắn     | Hành động, không khí | 図書館の扉を開けると、静寂が全身を包んだ。                                                            | Khi tôi mở cửa thư viện, sự im lặng bao trùm toàn bộ cơ thể tôi                                                                                                                                                                                | Đúng ý nhưng hơi cứng                                 |
| 3   | Trung    | Giáo dục, tâm lý     | 久しぶりの会話授業でしたが、口が重くなっていて、うまく話せないように感じましたね。やはり実際に日本語を話す機会がないと、どんどん下手になってしまふのでしょうか。 | Đã lâu rồi tôi mới tham gia cuộc trò chuyện, nhưng miệng tôi cảm thấy nặng trĩu và tôi cảm thấy như mình không thể diễn tả tốt được. Đúng như dự đoán, nếu bạn không có cơ hội thực sự nói chuyện với tiếng Nhật, bạn sẽ trở nên tệ hơn nhiều. | Dịch đúng ý, tự nhiên                                 |
| 4   | Trung    | Sinh hoạt, thư giãn  | 昨日は友達と近くの公園を散歩しました。春の陽気でとても気持ちよく、桜も少しずつ咲き始めていまし                                  | Hôm qua tôi đi dạo qua công viên gần đây với bạn bè. Thời tiết mùa xuân thật tuyệt vời, mang lại cảm giác dễ chịu, và những bông hoa anh đào bắt đầu nở rộ từng bông. Đã lâu rồi                                                               | Đúng ý nhưng còn lặp từ, cần gọn gàng và tự nhiên hơn |

|   |       |                              |                                                                                                                                                                |                                                                                                                                                                                                                                                                       |                                                                                                                                                                |
|---|-------|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
|   |       |                              | た。久しぶりにゆ<br>っくり話ができ<br>て、リフレッシュ<br>できました。                                                                                                                      | tôi mới có_thể<br>trò_chuyện thoải_mái<br>lần đầu_tiên sau một<br>thời_gian dài và đã có<br>một khoảng thời_gian<br>rảnh_rỗi .                                                                                                                                        |                                                                                                                                                                |
| 5 | Trung | Hành<br>động,<br>nội<br>tâm  | 半分ほど埋まっ<br>ていた駐輪場に自転<br>車を止めてから、<br>咲太は館内に足を<br>踏み入れた。何度<br>通っても図書館特<br>有の静けさには、<br>なかなか体が馴染<br>まない。わずかに<br>体が緊張する。                                            | Sakuta đỗ xe_đạp ở<br>bãi đỗ xe_đạp nơi<br>một_nửa thể_là đi vào<br>trong tòa nhà . Dù tôi<br>có đi qua bao_nhiều<br>lần , tôi vẫn không_thể<br>ngồi im trong<br>thư_viện , nơi thường<br>thấy yên_tĩnh<br>đặc_trung của<br>thư_viện . Cơ_thể tôi<br>hơi căng_thẳng . | Đúng ý<br>nhưng<br>dịch lặp<br>“bãi đỗ<br>xe_đạp”<br>và<br>“thư_viện<br>”, quá<br>literal, cần<br>gọn và tự<br>nhiên hơn.                                      |
| 6 | Trung | Thư<br>viện,<br>tạp<br>chí   | この辺では一番大<br>きい図書館だけ<br>に、利用者の数は<br>多い。入ってすぐ<br>のところにある雑<br>誌、新聞が置かれ<br>たコーナーには、<br>よく見かけるおじ<br>さんが今日もスポ<br>ーツ新聞を難しい<br>顔で読んでいた。<br>ひいきの球団が昨<br>日は負けたのだろ<br>うか。 | Có rất nhiều người<br>sử_dụng thư_viện lớn<br>nhất ở đây . Ở<br>khu_vực mà tôi vào có<br>tạp_chí và báo ngay<br>cạnh , và góc mà<br>người tôi thường thấy<br>đọc là đọc những tờ<br>báo thể_thao hôm_nay<br>với vẻ mặt phức_tạp .<br>Bạn có thua không ?              | Đúng ý<br>phần đầu<br>nhưng<br>dịch sai<br>ngữ cảnh<br>(“Bạn có<br>thua<br>không?”<br>không<br>khớp), cần<br>sửa lại câu<br>cuối cho<br>tự nhiên,<br>rõ nghĩa. |
| 7 | Trung | Siêu<br>nhiên,<br>cảm<br>xúc | 夜空に浮かぶ満月<br>を見上げて、僕は<br>小さく呟いた。                                                                                                                                | Nhìn lên bầu_trời đêm<br>đang lơ_lửng trên<br>bầu_trời đêm , tôi<br>lắm_bảm nhẹ_nhàng .<br>" Đây có thực_sự là                                                                                                                                                        | Đúng ý<br>nhưng<br>dịch lặp<br>“bầu_trời”<br>, dấu chấm                                                                                                        |

|   |       |                     |                                                                                                                                    |                                                                                                                                                                                                                                                                                                                              |                                                                                                                                             |
|---|-------|---------------------|------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
|   |       |                     | 「これが、本当に奇跡なんだろうか...」 風が耳元をかすめ、ひんやりとした予感が胸に残った。                                                                                     | một phép màu kỳ lạ không ? " Gió lướt qua tai tôi , thoáng qua trong tâm trí tôi .                                                                                                                                                                                                                                           | lững và ần dụ chưa tự nhiên, cần mượt câu hơn.                                                                                              |
| 8 | Ngắn  | Nội tâm, khung cảnh | 教室の隅で、彼女はノートを握り締めたまま固まっていた。その一言に、教室中の時間が止まった気がした。                                                                                  | Cô ấy đứng im ở góc lớp học , tay cầm một cuốn sổ tay . Chỉ một từ đó thôi đã khiến tôi cảm thấy như thời gian bên trong lớp học đã dừng lại .                                                                                                                                                                               | Đúng ý nhưng dịch còn literal, cấu trúc cứng và thiếu cảm giác “khoảnh khắc ngưng đọng” trong lớp học.                                      |
| 9 | Trung | Chiêm nghiệm, ký ức | 夕暮れの駅前広場には、わずかに赤い夕日が残っていた。彼は重いかばんを肩にかけ、ゆっくりと階段を上った。階段の上から見下ろすと、見慣れた景色がいつもより遠く感じられた。「変わったのは、世界じゃなくて、俺なんだろうな」。そんな思いとともに、彼の足は止まらなかった。 | Trong quảng trường trước nhà ga lúc chạng vạng , có một ánh hoàng hôn chiếu vàng . Anh ta đặt một tấm rèm nặng nề lên vai và từ từ leo lên cầu thang . Nhìn xuống từ đỉnh cầu thang , quang cảnh quen thuộc giờ đây có vẻ xa xôi hơn bình thường . " Sự thay đổi này không phải là thứ tôi cảm thấy như mình đã bị mắc kẹt . | Đúng ý cảnh nhưng “tám rèm” dịch sai “cặp”, dịch quá literal và mất mạch tâm trạng suy ngẫm; cần sửa chính xác chủ thể và làm câu mượt hơn. |

|    |     |                        |                                                                                                                                       |                                                                                                                                                                                                                                                                                           |                                                                                                                |
|----|-----|------------------------|---------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------|
|    |     |                        | た。                                                                                                                                    |                                                                                                                                                                                                                                                                                           |                                                                                                                |
| 10 | Dài | Môi trường, khoa học   | 地球温暖化が進行する中、海面上昇のペースはこれまでにない速さで加速している。専門家たちは、2050年までに世界の大規模沿岸都市の20%が浸水する可能性が高いと警告する。再生可能エネルギーの導入や二酸化炭素削減対策を急ぐ必要があるが、国際的な協力は依然として不十分だ。 | Khi sự toàn cầu tăng lên, tốc độ leo lên bề mặt của tàu vãn nhanh và nhanh hơn bình thường. Các chuyên gia cảnh báo rằng có khả năng cao là 20 thành phố ven biển lớn sẽ bị tác. Việc phân phối này sẽ dẫn đến việc tăng nguy cơ tổn thất lên hai hoặc ba tuyến đường khác.               | Mô hình hoàn toàn không hiểu văn bản chuyên ngành môi trường, dịch sai ngữ nghĩa trầm trọng.                   |
| 11 | Dài | Kinh tế vĩ mô, số liệu | 先週発表された統計によると、国内総生産（GDP）は前年度比3.2%増となり、予想を上回った。しかし、消費者物価指数（CPI）の上昇も同時に記録され、生活コストへの懸念が広がっている。政府は追加の景気刺激策として、低所得世帯への補助                   | Theo số liệu thống kê đã được công bố, [UNK] sản xuất cả nước sẽ là 32 và vượt xa mong đợi của hiện nay. Tuy nhiên, số tiền chi tăng / giảm được tạo thành từ [UNK], và mối lo ngại về chi phí ngày càng tăng. chính phủ này đã được ghi lại và sử dụng các biện pháp kích thích kinh tế. | Dịch sai nghiêm trọng, không hiểu thuật ngữ kinh tế, xuất hiện nhiều lỗi cú pháp và từ ngữ vô nghĩa như [UNK]. |

|    |     |                   |                                                                                                                                       |                                                                                                                                                                                                                                                                                                                         |                                                                                             |
|----|-----|-------------------|---------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|
|    |     |                   | 金拡充を検討中だ。                                                                                                                             |                                                                                                                                                                                                                                                                                                                         |                                                                                             |
| 12 | Dài | Hợp đồng, pháp lý | 本契約書は、発注者と受注者との間で締結される業務委託契約に関する一切の条件を定めるものとする。受注者は、納期および品質に関して発注者の指示に従い、契約期間内に成果物を提出する義務を負う。契約の解除および損害賠償については、本契約第7条および第8条に基づくものとする。 | Sách hợp đồng bao gồm một lệnh và một hợp đồng được quyết định dựa trên bậc phụ huynh và người đứng đầu. Những người theo hợp đồng yêu cầu nhân viên và số tiền này làm công tác kiểm nợ và các khoản nợ phải làm theo yêu cầu số 8, và số tài sản được nộp lại sẽ được nộp lên. Về mặt hợp đồng, vấn đề này là về việc | Dịch hoàn toàn sai nghĩa, không hiểu văn bản pháp lý, nội dung trở nên vô nghĩa và rời rạc. |

Dựa trên các ví dụ từ nhiều thể loại văn bản khác nhau, mô hình dịch thể hiện khả năng tương đối ổn định trong các văn bản thuộc thể loại light novel hoặc đời sống thường nhật, nhưng lại gặp nhiều lỗi nghiêm trọng với các văn bản chuyên ngành, học thuật, hoặc có cấu trúc ngôn ngữ phức tạp.

- **Ưu điểm – Dịch light novel, đời sống khá là tự nhiên.**

Với những câu mang sắc thái tự sự, miêu tả đời sống hoặc nội tâm nhân vật thường thấy trong light novel (ví dụ: câu 1, 2, 4, 5, 6, 7, 8), mô hình dịch đạt kết quả khá tốt. Dù vẫn còn tồn tại lỗi từ vựng hoặc cấu trúc chưa tự nhiên, bản dịch thường giữ được ý nghĩa chính, không sai lệch nhiều, và có thể hiệu chỉnh dễ dàng về mặt ngữ pháp hoặc phong cách. Đây là dấu hiệu cho thấy dữ liệu huấn luyện có lẽ được lấy phần lớn từ văn bản văn học hoặc hội thoại thông thường.

- **Hạn chế - Dịch chuyên ngành kém, sai lệch nghiêm trọng.**

Với các văn bản chuyên ngành như thời sự (câu 10), kinh tế (câu 11), hoặc pháp lý (câu 12), mô hình thể hiện rõ sự yếu kém: dịch sai khái niệm, lặp cấu trúc rồi

rầm, nhiều cụm từ vô nghĩa hoặc bỏ sót từ khóa quan trọng. Những lỗi này không chỉ gây mất nghĩa mà còn dẫn đến hiểu nhầm thông tin. Điều này chứng minh rằng mô hình không được huấn luyện đủ với dữ liệu có tính chuyên môn cao.

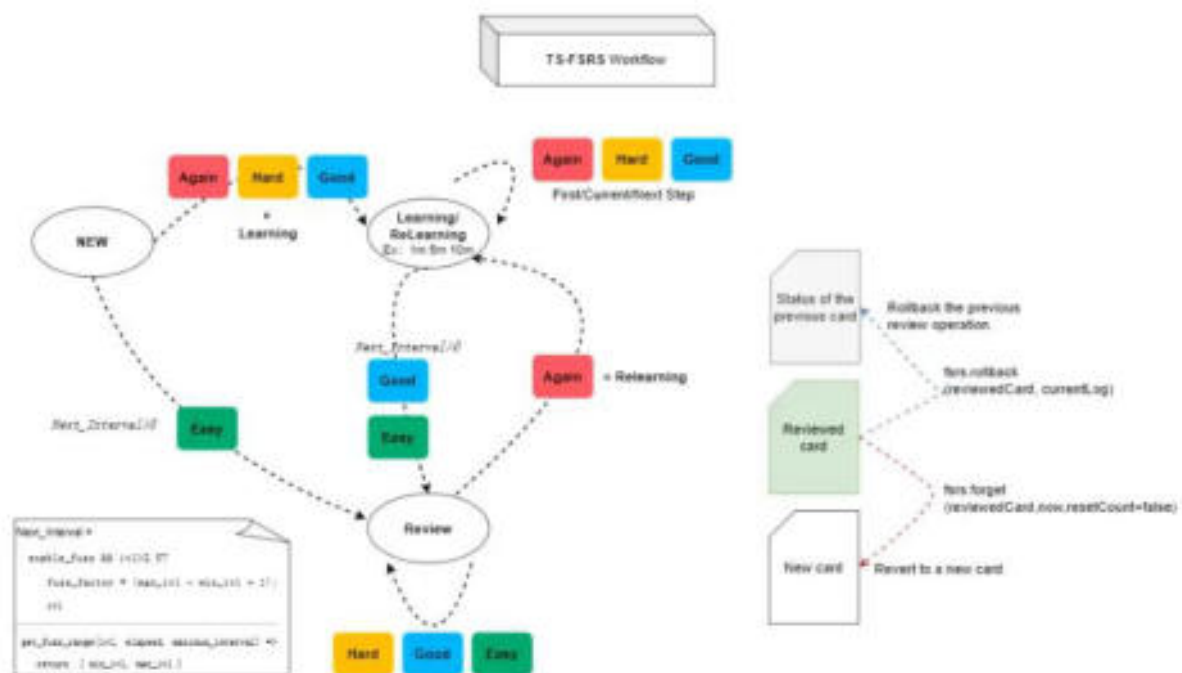
- Dịch câu dài yếu – giới hạn độ dài đầu vào.

Với những câu/đoạn dài (ví dụ: câu 5, 6, 9, 10, 11, 12), chất lượng dịch suy giảm rõ rệt. Mô hình bắt đầu mất mạch ngữ nghĩa, xáo trộn câu từ, thậm chí dẫn đến dịch lặp hoặc sai chủ ngữ – vị ngữ. Đây có thể là hậu quả trực tiếp từ giới hạn kiến trúc mô hình chỉ huấn luyện trên 64 token, khiến thông tin dài hạn không được duy trì đủ tốt trong quá trình sinh câu.

Tóm lại mô hình dịch hoạt động tốt với câu ngắn, văn phong gần gũi với hội thoại hoặc văn học nhẹ nhàng – đặc biệt là nội dung tương tự với light novel. Tuy nhiên, nó gặp khó khăn đáng kể khi xử lý các văn bản chuyên ngành, văn bản thông tin dài, hoặc có cấu trúc logic chặt chẽ. Nguyên nhân chủ yếu nằm ở thiếu dữ liệu chuyên ngành trong quá trình huấn luyện và giới hạn kiến trúc (64 token) khiến mô hình không nắm được ngữ cảnh dài.

#### 4.2.2. Triển khai thuật toán FSRS.

Để ứng dụng FSRS (Forgotten Spaced Repetition System) một cách hiệu quả, tôi đã xây dựng quy trình xử lý thẻ (card) như sau, dựa trên sơ đồ TS-FSRS Workflow:



Hình 36 Hoạt động sử dụng của thuật toán FSRS

### Khởi tạo thẻ mới (NEW)

- Khi người dùng thêm một từ vựng hoặc hán tự mới, hệ thống tạo một thẻ ở trạng thái **NEW**.

- Thẻ mới chưa có lịch ôn tập; ngay khi bắt đầu học lần đầu, thẻ được đẩy sang giai đoạn **Learning** với một chuỗi khoảng lặp ban đầu (ví dụ: 1 phút → 5 phút → 10 phút), tùy cấu hình của ứng dụng.

### Giai đoạn Learning / Re-learning

- **Mục đích:** Giúp người học làm quen ban đầu với từ mới, trước khi chuyển sang ôn định kỳ dài hạn.
- **Quy trình:**
  1. Hệ thống hiển thị thẻ và yêu cầu người dùng đánh giá mức độ ghi nhớ với ba lựa chọn:
    - **Again** (chưa nhớ)
    - **Hard** (khó nhớ)
    - **Good** (dễ nhớ)
  2. Nếu người dùng chọn **Again**, thẻ được đưa ngay về lại giai đoạn Learning (re-learning) để ôn lại ngay lập tức.
  3. Nếu chọn **Hard** hoặc **Good**, hệ thống sẽ tính toán Next\_Interval – khoảng thời gian đến lần ôn tập tiếp theo – dựa trên công thức:

```
if enable_fuzz && current_interval > min_interval:
    // Thêm độ nhiễu ngẫu nhiên nhằm tránh quá trình ôn tập quá "đồng đều"
    Next_Interval = fuzz_factor * (max_interval - min_interval + 1) + min_interval
else:
    // Tăng theo tỷ lệ cố định (ví dụ: x2 nếu Good, x1.2 nếu Hard)
    Next_Interval = current_interval * growth_factor
```

4. Khi Next\_Interval vượt ngưỡng min\_interval, thẻ được chuyển sang giai đoạn **Review**.

### Giai đoạn Review

- **Mục đích:** Ôn tập định kỳ ở mốc thời gian dài hơn, dựa vào mức ghi nhớ đã cải thiện.
- **Quy trình:**
  1. Đến thời điểm Next\_Interval, thẻ được đẩy vào lịch Review.
  2. Người dùng lại đánh giá **Again**, **Hard**, **Good**:
    - **Again** → thẻ quay trở lại Learning để ôn lập tức.

- **Hard/Good** → hệ thống tính tiếp Next\_Interval theo công thức tương tự, và lên lịch Review tiếp.
- Quá trình Review lặp lại cho đến khi người dùng đánh giá “Good” nhiều lần liên tiếp và Next\_Interval đủ dài để có thể cho là “thuộc”.

### Tùy chỉnh và cá nhân hóa

- **Tham số cấu hình:**
  - *min\_interval, max\_interval*: Giới hạn khoảng lặp nhỏ nhất và lớn nhất.
  - *fuzz\_factor*: Độ nhiễu ngẫu nhiên nhằm tránh hiệu ứng “đồng loạt ôn tập”.
  - *growth\_factor*: Tỷ lệ tăng khoảng lặp mỗi lần đánh giá “Good” hoặc “Hard”.
  - *enable\_fuzz*: Bật/tắt cơ chế thêm nhiễu.
- **Tối ưu hóa tự động:** Hệ thống thu thập dữ liệu ôn tập (thời gian hoàn thành, kết quả đánh giá...) để phân tích và gợi ý điều chỉnh tham số, từ đó cá nhân hóa lộ trình cho từng học viên.

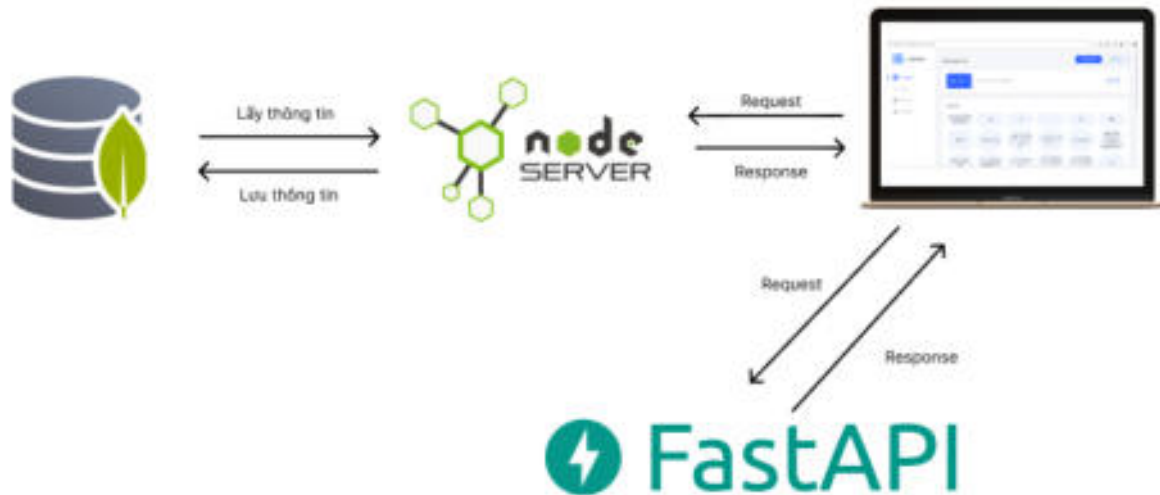
### Kết quả kỳ vọng

- **Ôn tập đúng lúc:** Tránh lãng phí thời gian cho từ đã thuộc và không dễ quên quá lâu dẫn đến ôn lại từ đầu.
- **Khả năng hoàn nguyên:** Người dùng an tâm thử nghiệm đánh giá, sai có thể hoàn tác mà không mất tiến trình.
- **Cá nhân hóa cao:** Mỗi người học có thể có lịch ôn tập riêng, tối ưu dựa trên khả năng ghi nhớ cá nhân.

Với triển khai này, FSRS được ứng dụng trọn vẹn từ giai đoạn làm quen ban đầu, ôn tập định kỳ đến quản lý lịch sử và tinh chỉnh tham số, mang lại trải nghiệm học thẻ hiệu quả, linh hoạt và chủ động cho người dùng.

## CHƯƠNG 5. TRIỂN KHAI THỰC TẾ

### 5.1. Cấu trúc hệ thống



Hình 37 Tổng quan về hệ thống

### 5.2. Môi trường phát triển

- Công cụ phát triển: Visual Studio Code, Postman
- Ngôn ngữ phát triển backend: Express.js
- Ngôn ngữ phát triển frontend: TypeScript, ReactJS
- Ngôn ngữ phát triển AI: Python, FastAPI
- Cơ sở dữ liệu: MongoDB
- Quản lý mã nguồn: Git, Github
- Môi trường để chạy chương trình:
  - Node.js: Phiên bản v22.14.0
  - Python: Phiên bản v3.11.12
  - MongoDB: Phiên bản v8.0

### 5.3. Môi trường triển khai

- Cơ sở dữ liệu: MongoDB Atlas
- Môi trường deploy: EC2 (AWS)

#### Mục tiêu:

Triển khai toàn bộ dịch vụ Frontend (React + Vite), Backend (Node.js/Express) lên một

EC2 duy nhất, đồng thời giữ AI service (FastAPI) chạy trên máy local để tiết kiệm tài nguyên EC2.

**Kiến trúc hệ thống:**

- EC2 khởi chạy Docker Compose với hai container:
  - frontend: serve static files build từ React/Vite thông qua Nginx (port 80).
  - backend: cung cấp REST API (port 8080).
- AI Service chạy trên máy local, được kết nối với EC2 qua SSH reverse tunnel (mở port nội bộ EC2, ví dụ 9001) để frontend/backend có thể gọi API mà không cần deploy model lên cloud.

**Luồng dữ liệu:**

- Người dùng truy cập `http://<ElasticIP>/` → Nginx tại frontend container.
- Các request `/api/v1/...` được proxy về backend container.
- Các request `/fastapi/api/v1/...` (REST) và WebSocket `/ws/translate` được proxy qua tunnel (`127.0.0.1:9001`) tới AI server trên local.
- Kết quả dịch thuật trả ngược qua tunnel về EC2, Nginx chuyển tiếp về trình duyệt.

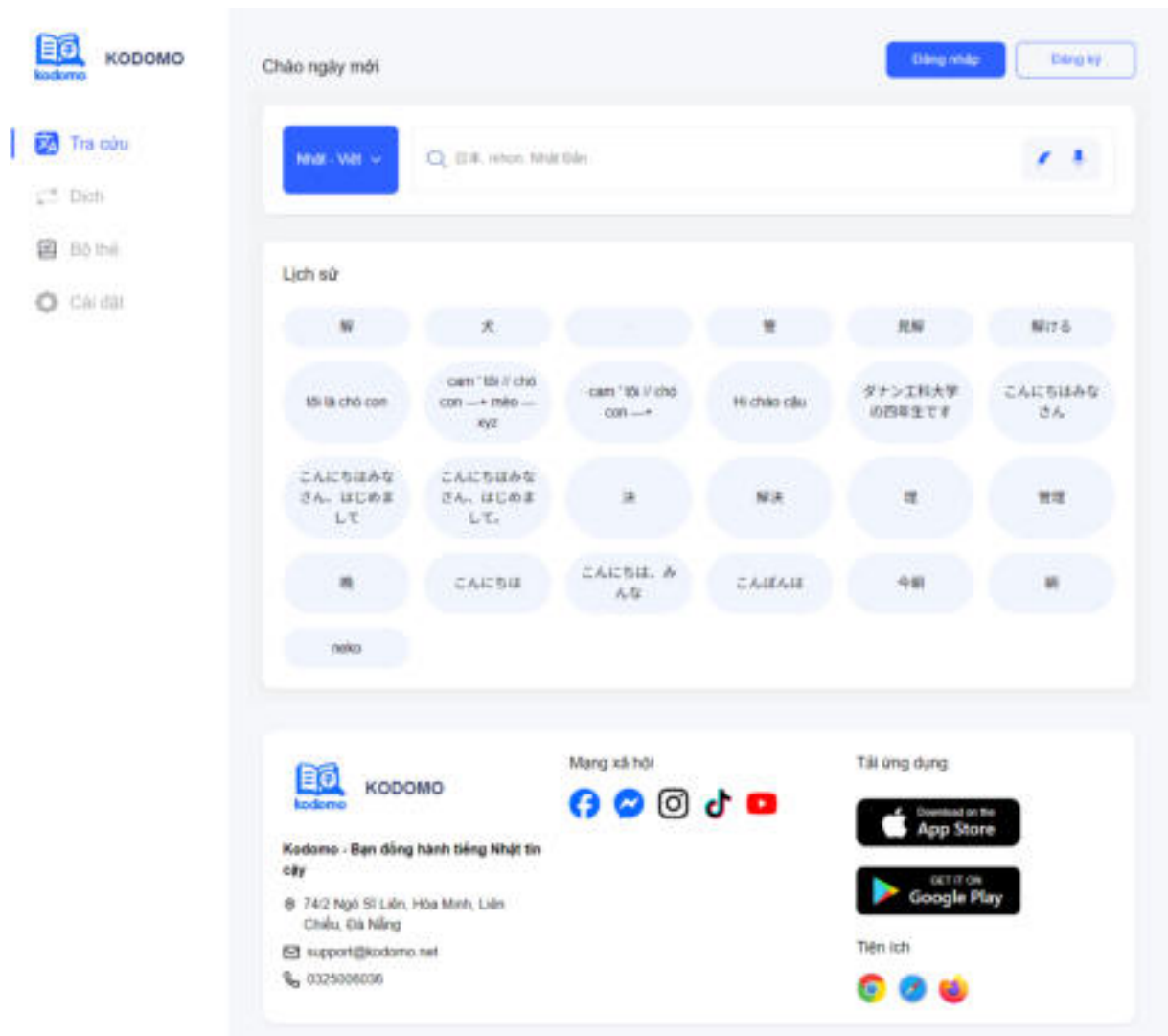
**5.4. Kết quả thực tế**

Hệ thống này bao gồm 2 view chính:

- View người dùng.
- View khách.

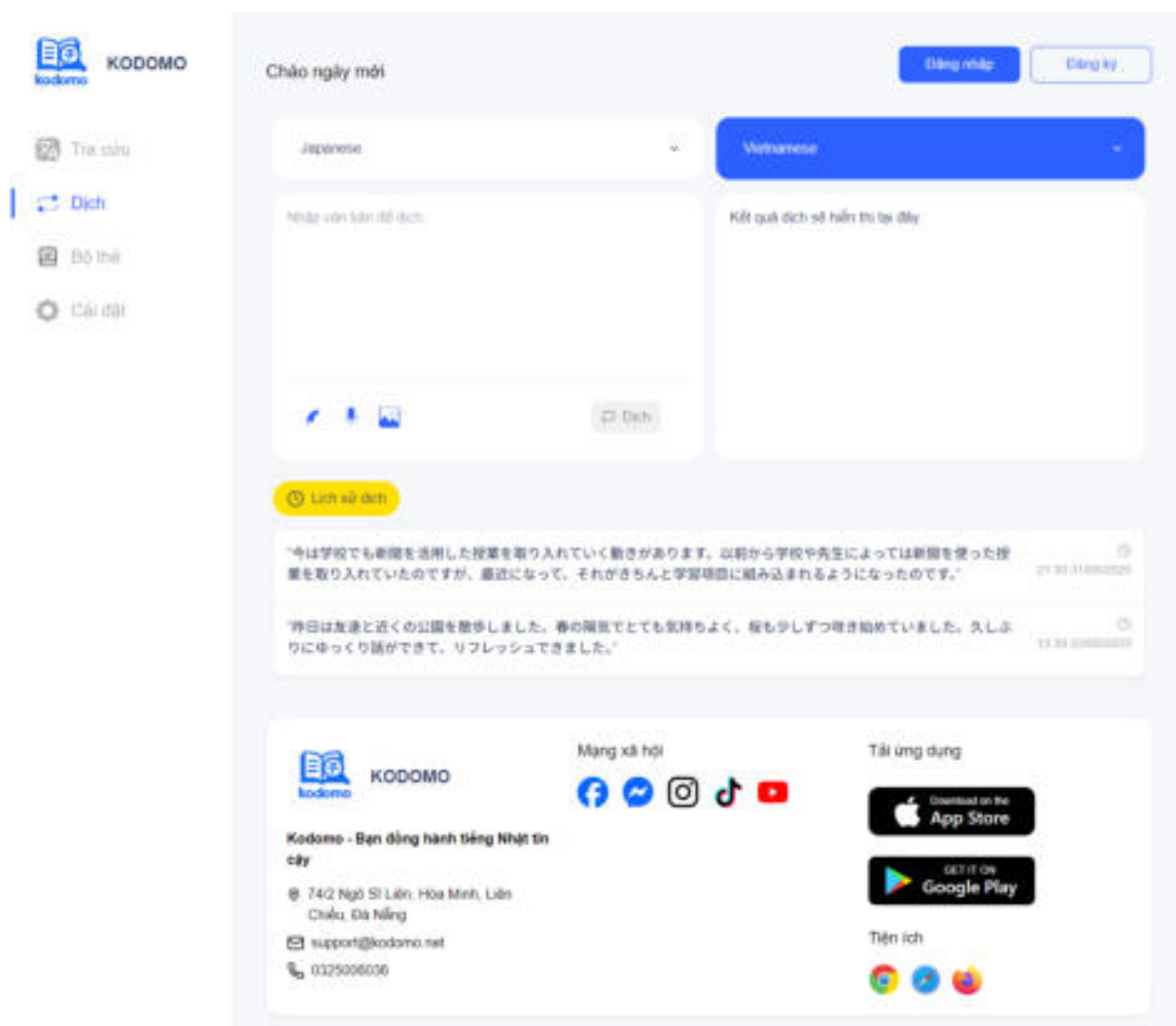
**5.4.1. Giao diện người dùng**

**Giao diện tra cứu từ vựng:** Giao diện trang chủ tối giản, tập trung vào chức năng chính là tra cứu từ vựng với các nút trợ năng ngay bên cạnh thanh tìm kiếm. Ngoài ra còn hiển thị danh sách lịch sử giúp người dùng dễ dàng tìm lại những từ đã tra cứu.



Hình 38 Giao diện tra cứu từ vựng

**Giao diện dịch:** Dịch câu văn từ tiếng Nhật sang tiếng Việt và ngược lại. Có các nút trợ năng như nút OCR hình ảnh ra chữ, nút mở board vẽ và detect ra chữ kanji, nút ghi âm giúp ghi âm giọng nói ra chữ. Các nút trợ năng giúp người dùng dễ dàng truyền tải câu văn muốn dịch hơn.



Hình 39 Giao diện dịch

**Giao diện đăng nhập:** Giao diện đăng nhập thiết kế đơn giản chỉ cần nhập 2 trường là email và password là có thể đăng nhập và sử dụng.

KODOMO

Chào ngày mới

Đăng ký

Đăng nhập với

Google

Email

Enter your email

Mật khẩu

Enter your password

Quên mật khẩu

Đăng nhập

Bạn chưa có tài khoản? Đăng ký

KODOMO

Mạng xã hội

Tải ứng dụng

Download on the App Store

GET IT ON Google Play

Tiện ích

Kodomo - Bạn đồng hành tiếng Nhật tin cậy

@ 74/2 Ngõ Sĩ Lân, Hòa Minh, Liên Châu, Đà Nẵng

support@kodomo.net

0325906006

Hình 40 Giao diện đăng nhập

**Giao diện đăng ký:** Giao diện đăng ký thiết kế đơn giản, chỉ cần nhập email và mật khẩu là đã có thể đăng ký được tài khoản. Giúp người dùng đăng ký được tài khoản nhanh chóng.

KODOMO

Chào ngày mới

Đăng nhập

Đăng ký với

Google

Email

Enter your email

Mật khẩu

Enter your password

Nhập lại mật khẩu

Re-enter your password

Đăng ký

KODOMO

Mạng xã hội

Tải ứng dụng

Download on the App Store

GET IT ON Google Play

Tiện ích

Kodomo - Bạn đồng hành tiếng Nhật tin cậy

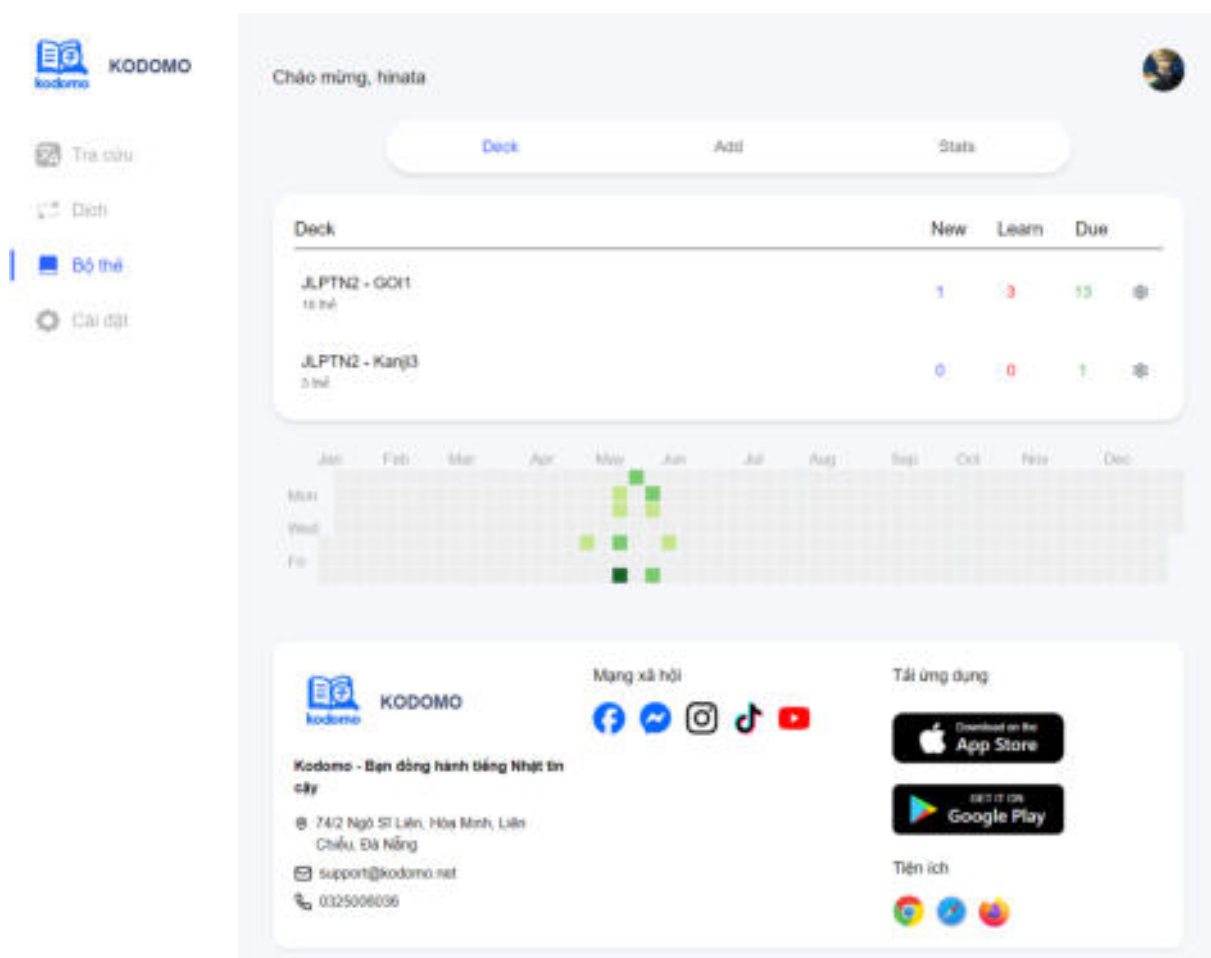
74/2 Ngõ Sĩ Liên, Hòa Minh, Liên Châu, Đà Nẵng

support@kodomo.net

0325006006

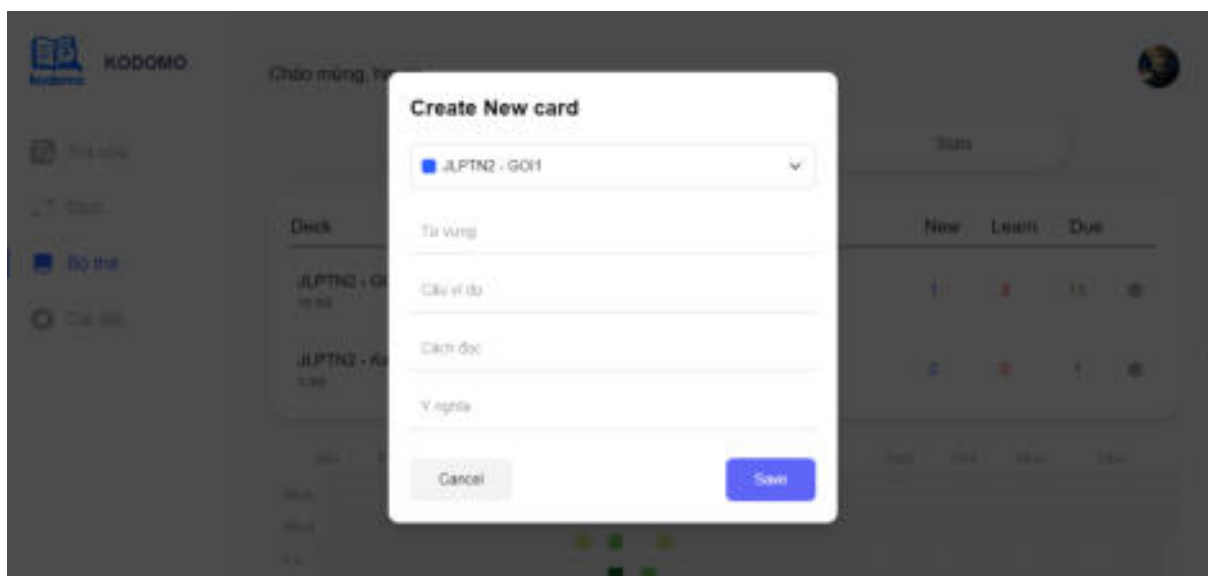
Hình 41 Giao diện đăng ký

**Giao diện bộ thẻ:** Hiển thị danh sách các bộ thẻ của người dùng. Trên mỗi bộ thẻ sẽ hiển thị các thông tin như: tên bộ thẻ, tổng số thẻ trong bộ thẻ, số thẻ mới, số thẻ đang học, số thẻ đến hạn học trong bộ thẻ. Ngoài ra còn có biểu đồ heatmap để thống kê tiến trình học của người dùng. Bên cạnh mỗi bộ thẻ có icon bánh răng dùng để cập nhật tên bộ thẻ, thêm thẻ vào bộ thẻ, xóa thẻ.



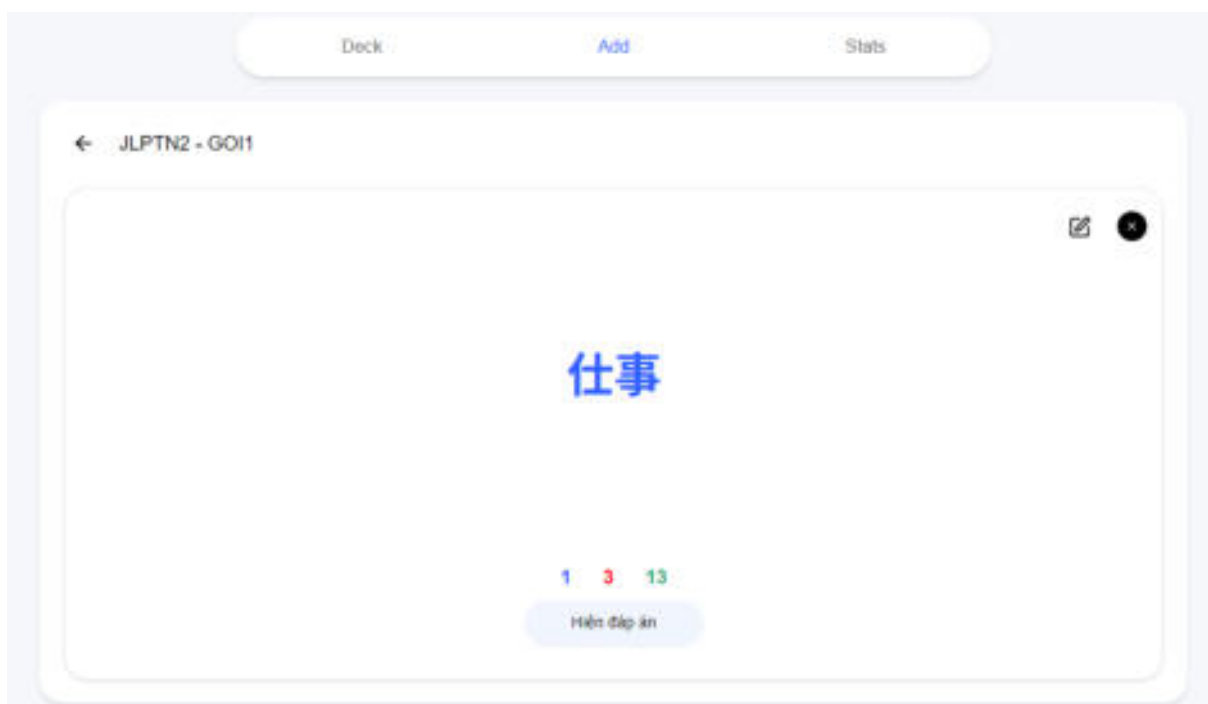
Hình 42 Giao diện bộ thẻ

**Giao diện modal thêm thẻ vào bộ thẻ:** Hiển thị các trường thêm thẻ mới vào bộ thẻ như: Từ vựng, Câu ví dụ, Cách đọc, Ý nghĩa. Ngoài ra còn hiển thị danh sách các bộ thẻ hiện có để người dùng tùy chỉnh thêm thẻ vào bộ thẻ tùy ý



Hình 43 Giao diện modal thêm thẻ vào bộ thẻ

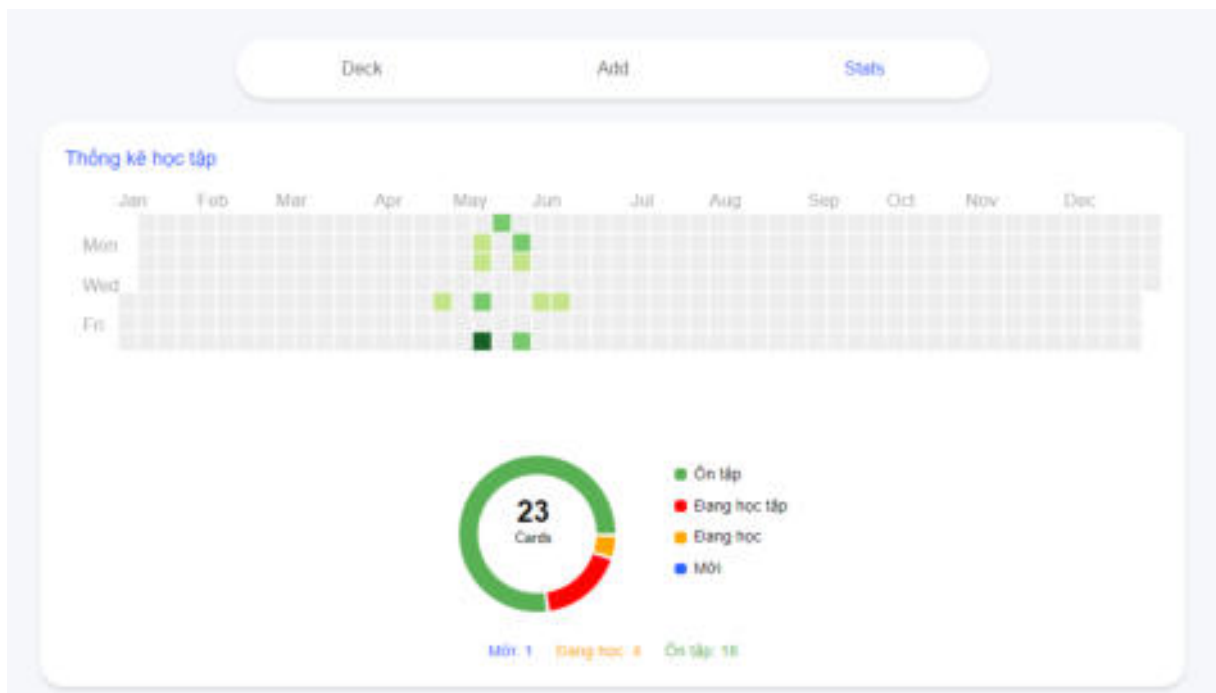
**Giao diện học thẻ:** Hiển thị học thẻ theo kiểu flashcard. Khi học thẻ người dùng sẽ tùy thuộc vào khả năng ghi nhớ từ của mình để chọn các action: Again, Hard, Good, Easy. Hệ thống sẽ sử dụng thuật toán FSRs để tính toán trạng thái thẻ học cho người dùng.





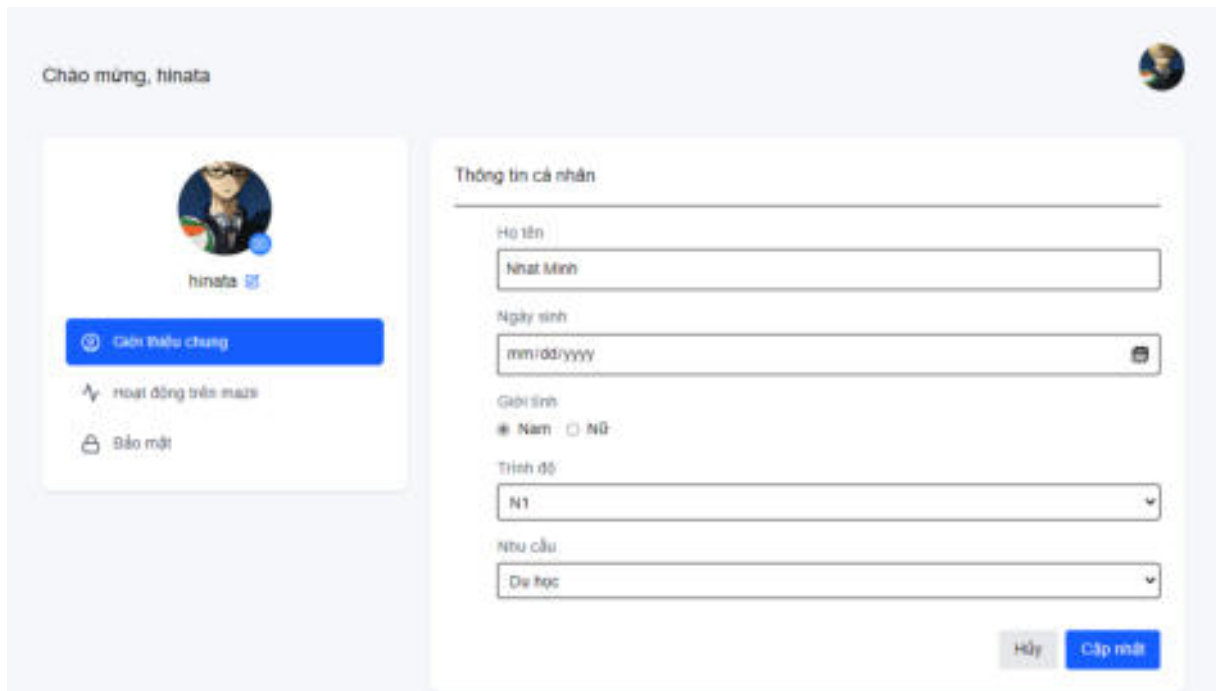
Hình 44 Giao diện học thẻ

**Giao diện thống kê thẻ học:** Hiện thị biểu đồ heatmap để thể hiện tiến trình học thẻ của người dùng và biểu đồ hình tròn thể hiện trạng thái thẻ học cho người dùng.



Hình 45 Giao diện thống kê thẻ học

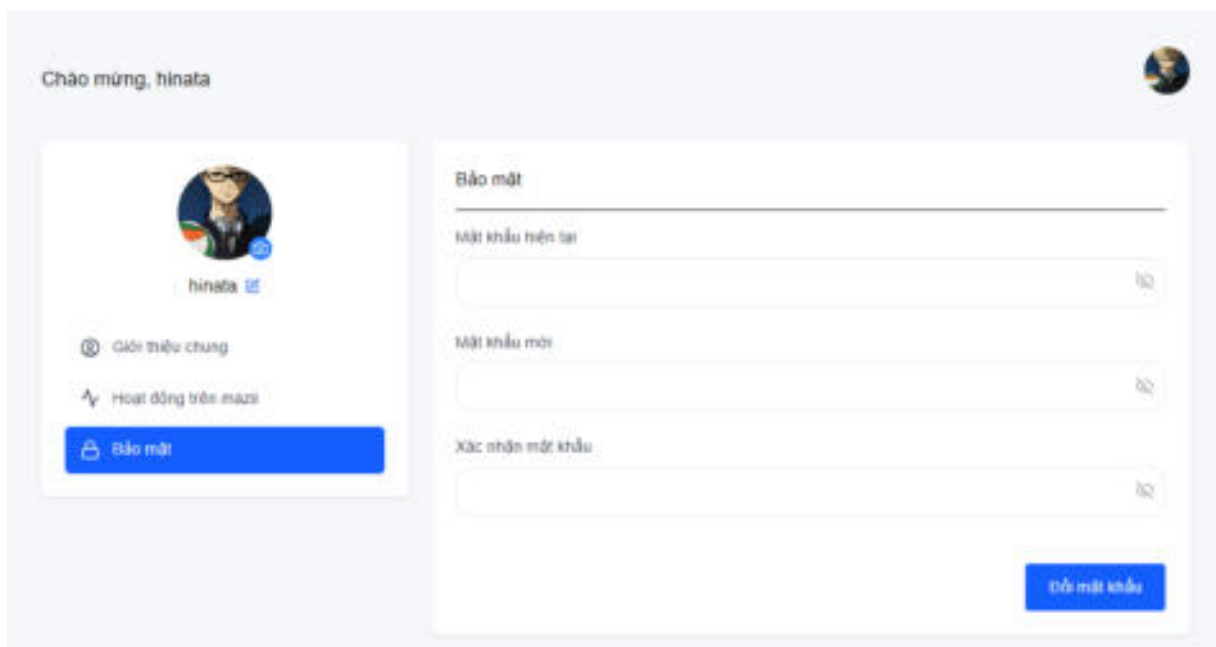
**Giao diện cập nhật thông tin người dùng:** Hiển thị các thông tin cơ bản người dùng và người dùng có thể dễ dàng cập nhật thông tin cá nhân của mình.



The screenshot shows a user profile update page. On the left, there's a sidebar with a user profile picture and name 'hinata', and three menu items: 'Giới thiệu chung', 'Hoạt động trên mạng', and 'Bảo mật'. The main content area is titled 'Thông tin cá nhân' and contains several form fields: 'Họ tên' (Last name) with the value 'Nhật Minh', 'Ngày sinh' (Date of birth) with the value 'mm/dd/yyyy', 'Giới tính' (Gender) with radio buttons for 'Nam' (selected) and 'Nữ', 'Trình độ' (Education level) with a dropdown menu showing 'N1', and 'Nhu cầu' (Requirements) with a dropdown menu showing 'Đại học'. At the bottom right, there are two buttons: 'Hủy' (Cancel) and 'Cập nhật' (Update).

Hình 46 Giao diện cập nhật thông tin người dùng

**Giao diện cập nhật mật khẩu:** Hiển thị form cập nhật mật khẩu đơn giản.

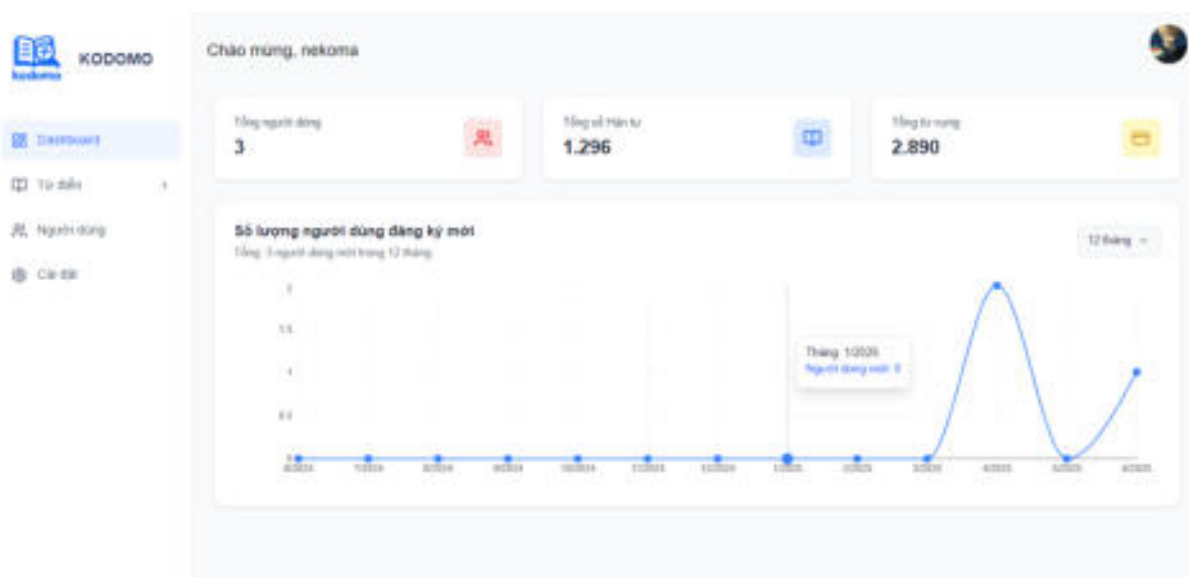


The screenshot shows a password update page. On the left, there's a sidebar with a user profile picture and name 'hinata', and three menu items: 'Giới thiệu chung', 'Hoạt động trên mạng', and 'Bảo mật'. The main content area is titled 'Bảo mật' and contains three form fields: 'Mật khẩu hiện tại' (Current password), 'Mật khẩu mới' (New password), and 'Xác nhận mật khẩu' (Confirm password). Each field has a toggle icon on the right. At the bottom right, there is a blue button labeled 'Đổi mật khẩu' (Change password).

Hình 47 Giao diện cập nhật mật khẩu

### 5.4.2. Giao diện quản trị viên

**Giao diện thống kê hệ thống:** Hiển thị thông tin thống kê về số người dùng, tổng số hán tự, tổng số từ vựng và biểu đồ thể hiện số người dùng đăng ký theo thời gian.



Hình 48 Giao diện thống kê hệ thống

**Giao diện danh sách từ vựng:** Hiển thị danh sách từ vựng phân trang.

The screenshot displays the KODOMO admin dashboard with the 'Danh sách từ vựng' (Vocabulary list) section active. The table below shows the list of vocabulary items:

| TỪ VỰNG | HÁN GIỮA | ROMAJI | Từ điển                                                                                                        | Thêm từ vựng |
|---------|----------|--------|----------------------------------------------------------------------------------------------------------------|--------------|
| A       | ア        | ha     | Người, anh người, nhân loại, vật, người được ưa, người có tài, người không thành người khác, hành thành, thành | →            |
| BB      | BB       | ashita | Ngày mai                                                                                                       | →            |
| BB      | BB       | are    | Cái là                                                                                                         | →            |
| BB      | BB       | are    | Cái là                                                                                                         | →            |
| B       | ウ        | hi     | Ngày, thời gian trong ngày, mặt trời, ánh sáng, mặt trời                                                       | →            |
| CC      | CC       | isaku  | Chỉ này                                                                                                        | →            |
| BB-CC   | BB-CC    | asaka  | Chỉ kia, đồng kia                                                                                              | →            |
| BB      | BB       | isaku  | Chỉ đi                                                                                                         | →            |
| CC-G    | CC-G     | isaku  | Phía này                                                                                                       | →            |
| BB-G    | BB-G     | isaku  | Phía đi                                                                                                        | →            |

At the bottom of the table, it says 'Tổng 2,000 bản ghi. Số bản ghi 1 trang' and 'Trang 1 / 200'.

Hình 49 Giao diện danh sách từ vựng

**Giao diện modal thêm và cập nhật từ vựng:** Hiển thị modal với các thông tin đầy đủ để có thể dễ dàng thêm và cập nhật từ vựng

**Cập nhật từ vựng** ✕

Điền thông tin từ vựng. Vui lòng bấm lưu xong khi điền xong.

Từ vựng:

Cách đọc:  + Thêm cách đọc

Ý nghĩa: Danh từ ▼

Nội dung ý nghĩa:

**Ví dụ 1** ✕

Câu:

Hiragana:

**Ví dụ 2** ✕

Câu:

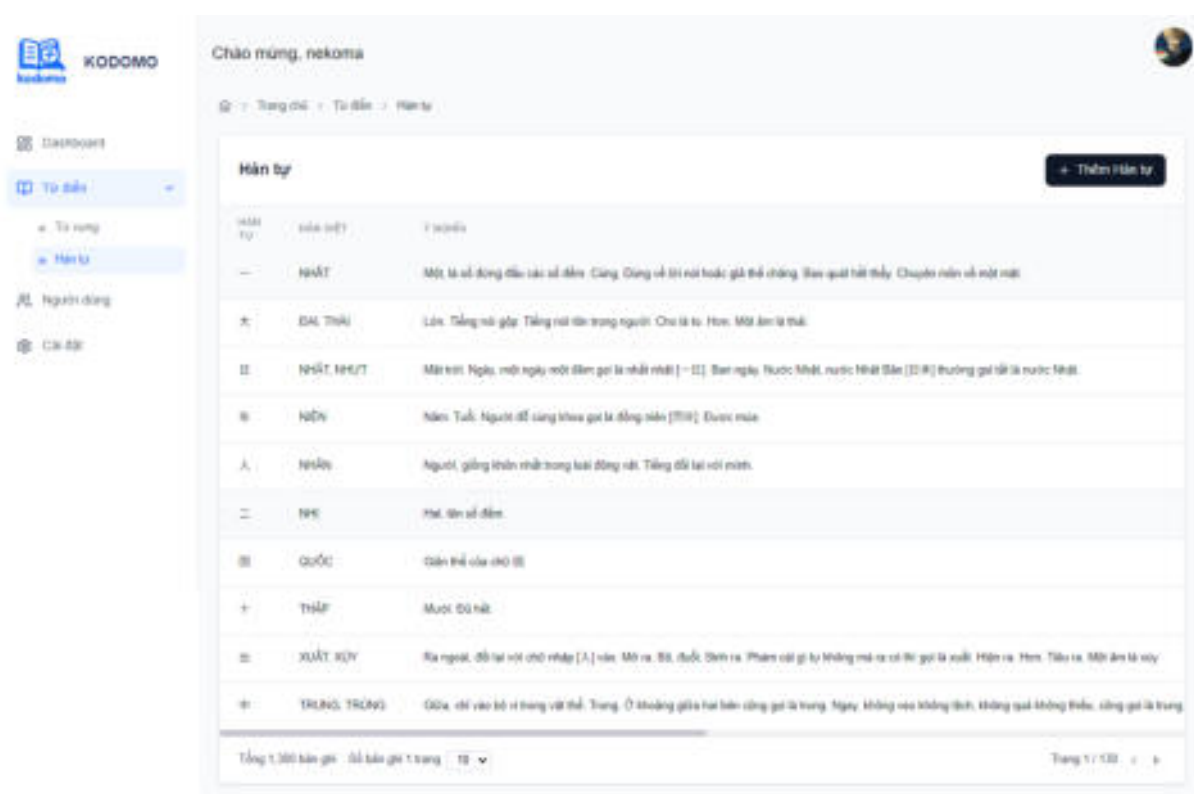
Hiragana:

Ý nghĩa:

+ Thêm ý nghĩa

Lưu

Hình 50 Giao diện modal thêm và cập nhật từ vựng

**Giao diện danh sách hán tự:** Hiển thị danh sách hán tự phân trang.

Hình 51 Giao diện danh sách hán tự

**Giao diện modal thêm và cập nhật hán tự:** Hiển thị modal với các thông tin đầy đủ để có thể dễ dàng thêm và cập nhật hán tự

### Cập nhật Hán tự

Điền thông tin Hán tự. Vui lòng bấm lưu xong khi điền xong.

Từ vựng

JLPT

N5

Số nét

Bộ thủ

Ý nghĩa

Hán Việt

Âm on

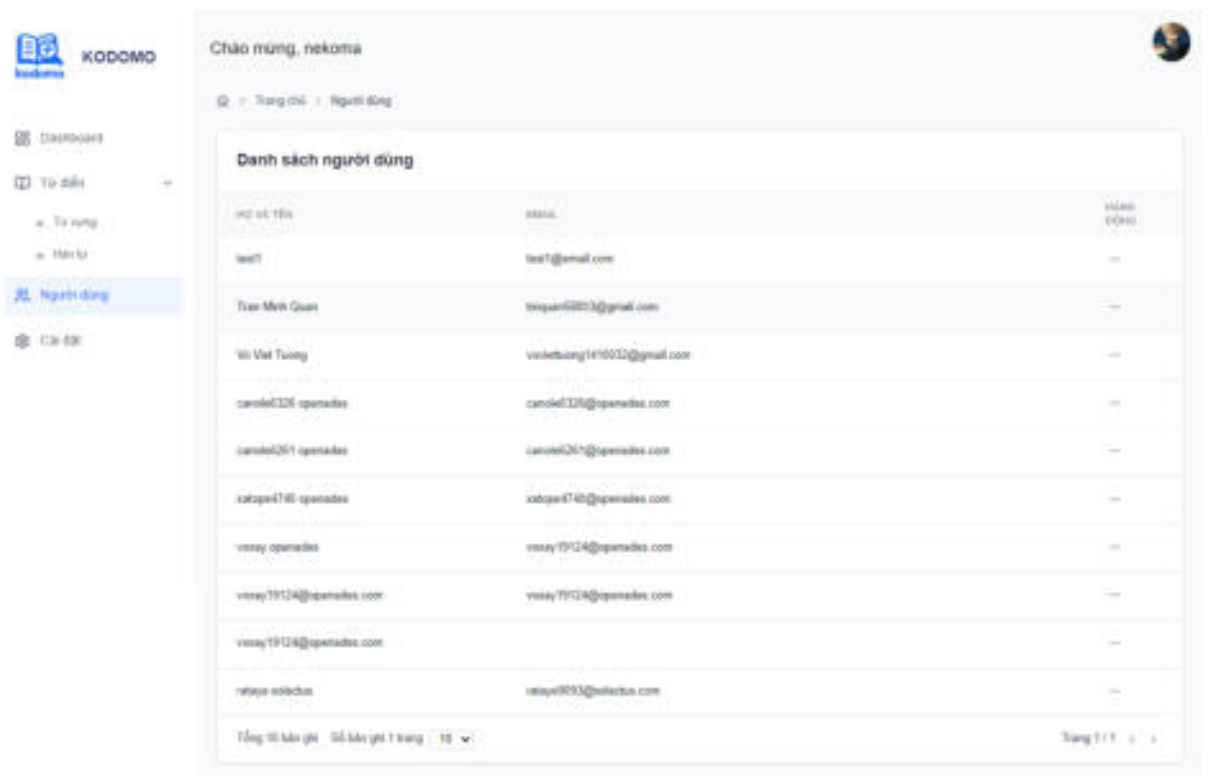
Âm kun

Hủy

Lưu

Hình 52 Giao diện modal thêm và cập nhật hán tự

## Giao diện danh sách người dùng: Hiển thị các thông tin cơ bản của người dùng



Hình 53 Giao diện danh sách người dùng

## KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

### 1.1. Kết quả đạt được

Trong quá trình nghiên cứu, tìm hiểu lý thuyết và thực hiện dự án, dự án đã thành công nhất định với mục tiêu ban đầu.

#### Về mặt lý thuyết:

- Trau dồi được kỹ năng phân tích, thiết kế và xây dựng cơ sở dữ liệu, bao gồm việc xác định các yêu cầu, thiết kế cấu trúc dữ liệu, và triển khai cơ sở dữ liệu cho hệ thống.
- Hiểu và vận dụng được ngôn ngữ lập trình Typescript, Python, các framework, thư viện như ExpressJS, ReactJS, FastAPI để xây dựng được hệ thống.
- Nghiên cứu và hiểu được cơ bản cách các mô hình hiện nay xử lý ngôn ngữ và áp dụng được vào trong một trường hợp cụ thể là dịch thuật

#### Về mặt thực tiễn ứng dụng:

- Trang web đáp ứng đầy đủ chức năng đề xuất ban đầu
- Trang web thuận tiện và linh hoạt cho người học, cho phép người học tra cứu từ vựng, Hán tự nhanh chóng. Ngoài ra còn sử dụng AI dịch thuật giúp cung cấp cho người dùng những bản dịch đúng, sát nghĩa và hợp với ngữ cảnh.
- Dễ dàng học tập và ghi nhớ từ vựng khi hệ thống áp dụng thuật toán FSRS (thuật toán ghi nhớ ngắt quãng) vào việc học từ.

### 1.2. Những khó khăn và cách khắc phục

Trong quá trình phát triển hệ thống, em đã gặp phải một số khó khăn và đã có những phương án khắc phục như sau:

- Trong quá trình training model dịch thuật sử dụng mô hình Transformer đã gặp một số khó khăn nhất định như:
  - Mô hình được xây dựng dựa trên tensorflow version 2.15.0 nên cần phải được training dựa trên bản tensorflow này để đạt được kết quả training tốt nhất. Nhưng hiện tại việc training trên google colab hay trên kaggle thì đều không hỗ trợ GPU cho tensorflow version 2.15.0. Điều này đã khiến em phải nâng version của tensorflow lên version 2.18.0 để có hỗ trợ GPU cho quá trình training. Việc nâng phiên bản lên như này có một hạn chế là ở phiên bản 2.18.0 không có hỗ trợ cơ chế mask cho transformer (Cơ chế mask là một cơ chế quan trọng để phân biệt giữa token thực và token đệm (padding) để Transformer học đúng và nhanh hơn trên các chuỗi có độ dài biến thiên, đồng thời giữ nguyên tính chất autoregressive của quá trình

sinh.)

*Giải pháp: Tự điều chỉnh mô hình transformer để bật cơ chế mask và lan truyền qua từng bước.*

- Vì trong quá trình training có sử dụng một lượng lớn data để training nhằm đạt được kết quả tốt nhất nhưng với lượng data lớn cũng đồng nghĩa cần phải có tài nguyên training đủ mạnh để training model. Trong quá trình training mặc dù đã đăng ký google colab pro để sử dụng nhưng vẫn không thể training với một lượng dữ liệu quá lớn.

*Giải pháp: Chia nhỏ dữ liệu ra training từng phần một. Đồng thời giới hạn số token training là 64 token. Hơn nữa cần phải điều chỉnh kích thước của mô hình xuống mức tối thiểu. Điều này sẽ giảm độ chính xác của kết quả đầu ra nhưng với một lượng data lớn thì điều này là cần thiết.*

- Để đạt được kết quả training tốt nhất thì cần phải có một bộ data lớn và sạch. Data được chuẩn bị từ các nguồn tiểu thuyết tiếng Nhật. Vì vậy trước khi đưa vào training cần phải chuẩn bị một bộ data dịch sang tiếng Việt chính xác. Ngoài ra ở mỗi dòng data sẽ có nhiều câu rất dài, trùng lặp và có chứa các ký tự không cần thiết vì vậy để đảm bảo data sạch đưa vào training cần phải sửa đổi rất nhiều trên một lượng lớn data.

*Giải pháp: Sử dụng và viết các file python để xử lý và làm sạch data. Ngoài ra cần phải giám sát kiểm tra thủ công trên vài mẫu data để đảm bảo lệnh được viết bởi python đã xử lý đúng.*

- Thiếu kinh nghiệm trong việc triển khai dự án lên môi trường AWS dẫn đến gặp khó khăn trong việc cấu hình, bảo mật và tối ưu hóa hiệu suất.

*Giải pháp: Nghiên cứu và áp dụng các công cụ hỗ trợ triển khai như Docker và AWS.*

- Vì thời gian phát triển có hạn nên website vẫn còn rất nhiều hạn chế và cần có nhiều cải tiến như cải tiến mô hình dịch thuật trên một bộ data lớn hơn, hay là tích hợp khả năng người dùng tự custom trong việc tạo thẻ học,...

*Giải pháp: Ưu tiên phát triển các chức năng cốt lõi nhằm đảm bảo hệ thống có thể vận hành ổn định. Các chức năng chưa hoàn thiện được ghi nhận để tiếp tục phát triển trong giai đoạn sau.*

### 1.3. Những vấn đề còn tồn tại

- Quản trị viên: Do hạn chế về thời gian, một số chức năng quản trị viên chưa được triển khai đầy đủ.
- Hiệu suất: Khi có nhiều người truy cập cùng lúc, trang web có thể chậm.
- Giao diện: Giao diện hiện tại cần cải thiện để trở nên đẹp mắt và thân thiện hơn.
- Tính năng và nội dung: Hệ thống vẫn chưa thu thập đủ nguồn tài liệu về bộ từ điển Việt – Nhật, vẫn còn những từ còn thiếu. Về việc học thẻ ghi nhớ, hiện tại người dùng chỉ mới có thể thêm thẻ học với những thông tin cơ bản như ý nghĩa, cách

đọc, câu ví dụ. Cần mở rộng để người dùng có thể tùy ý thêm những thông tin cần thiết về từ vựng phù hợp với phong cách học của mỗi người.

- Mô hình AI dịch thuật: Với lượng data training và tài nguyên training còn hạn chế cùng với việc mô hình chưa thực sự tối ưu đã làm cho mô hình AI dịch thuật vẫn chưa đáp ứng được kết quả như yêu cầu. Hiện tại mô hình AI vẫn chỉ dịch được những câu ngắn và văn phong ngữ nghĩa trong bản dịch có những đoạn vẫn chưa hợp lý.

#### **1.4. Hướng phát triển**

Mặc dù dự án từ điển Việt - Nhật đã đạt được nhiều kết quả tích cực, vẫn còn những hạn chế cần khắc phục để hoàn thiện hệ thống và mang lại trải nghiệm tốt hơn cho người dùng. Dựa trên các vấn đề đã được nhận diện, các hướng phát triển trong tương lai bao gồm:

##### **1. Mở rộng cơ sở dữ liệu từ điển**

- Tiếp tục thu thập, làm sạch và bổ sung dữ liệu song ngữ Việt - Nhật nhằm tăng số lượng và chất lượng mục từ trong từ điển.
- Hợp tác với cộng đồng người học và người dùng để bổ sung dữ liệu từ vựng thực tế và các ngữ cảnh sử dụng phổ biến.

##### **2. Cải thiện mô hình Transformer**

- Nâng cấp phần cứng và sử dụng các nền tảng điện toán đám mây mạnh mẽ hơn để huấn luyện mô hình trên các tập dữ liệu lớn hơn và đa dạng hơn.
- Thu thập thêm dữ liệu dịch thuật có độ dài câu khác nhau nhằm giảm giới hạn của mô hình khi xử lý câu dài.

##### **3. Đa dạng hóa chức năng học thẻ**

- Tích hợp thêm các loại thẻ học mới như thẻ hình ảnh, thẻ ví dụ ngữ cảnh, và thẻ kiểm tra âm thanh để tăng tính tương tác và hỗ trợ nhiều kiểu học tập khác nhau.
- Phát triển tính năng tự động đề xuất thẻ dựa trên lịch sử học tập và sở thích của người dùng.

##### **4. Cải thiện tính năng trích xuất chữ từ ảnh**

- Sử dụng các mô hình OCR tiên tiến hơn và tối ưu hóa thuật toán để tăng độ chính xác khi nhận diện chữ từ hình ảnh.
- Bổ sung các bộ dữ liệu hình ảnh đa dạng và phong phú hơn để mô hình học được các dạng chữ viết khác nhau.

##### **5. Phát triển ứng dụng di động**

- Để người dùng dễ dàng tra cứu, học thẻ ghi nhớ và sử dụng ứng dụng thì việc phát triển một ứng dụng di động là cần thiết.

## **6. Xây dựng hệ thống đánh giá và phản hồi**

- Phát triển chức năng bình luận và đánh giá nội dung trong từ điển.
- Tích hợp hệ thống gợi ý từ và chỉnh sửa nội dung từ các ý kiến đóng góp, tạo cơ chế kiểm duyệt để đảm bảo chất lượng.

## **7. Nghiên cứu và tích hợp công nghệ mới**

- Xem xét áp dụng các công nghệ tiên tiến như GPT, BERT, hoặc các mô hình xử lý ngôn ngữ tự nhiên mới để nâng cao khả năng dịch thuật và gợi ý nội dung.
- Phát triển thêm các tính năng như chatbot hỗ trợ học ngôn ngữ hoặc bài kiểm tra theo chủ đề.
- Những định hướng trên không chỉ nhằm khắc phục các hạn chế hiện tại mà còn mở rộng tiềm năng phát triển lâu dài của dự án, đáp ứng ngày càng tốt hơn nhu cầu học tập và sử dụng tiếng Nhật của người dùng.

## TÀI LIỆU THAM KHẢO

- [1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30. Retrieved from <https://arxiv.org/abs/1706.03762>
- [2] TensorFlow. (n.d.). Transformer model for language understanding. *TensorFlow Text Tutorials*. Retrieved April 22, 2025, from <https://www.tensorflow.org/text/tutorials/transformer>
- [3] TensorFlow. (n.d.). Subwords tokenizer. *TensorFlow Text Guide*. Retrieved April 29, 2025, from [https://www.tensorflow.org/text/guide/subwords\\_tokenizer](https://www.tensorflow.org/text/guide/subwords_tokenizer)
- [4] Meta. (n.d.). *React Documentation*. Retrieved April 01, 2025, from <https://react.dev/learn>
- [5] Express.js. (n.d.). *Express Documentation*. Retrieved April 03, 2025, from <https://expressjs.com>
- [6] Ramírez, S. (n.d.). *FastAPI Documentation*. Retrieved April 10, 2025, from <https://flask.palletsprojects.com>
- [7] Liu, H. (2022). FSRS: A modern algorithm for efficient spaced repetition. Retrieved April 20, 2025, from <https://github.com/open-spaced-repetition/fsrs4anki>
- [8] CaptainDario. (n.d.). *DaKanji-Single-Kanji-Recognition: A machine learning model to recognize Japanese characters (Kanji, Katakana, Hiragana)*. Retrieved April 12, 2025, from <https://github.com/CaptainDario/DaKanji-Single-Kanji-Recognition>