

**BÁO CÁO DỰ ÁN MÔN HỌC**  
**ĐỀ TÀI: XÂY DỰNG CHATBOT DỰA TRÊN LUẬT HỖ TRỢ TRA CỨU**  
**THÔNG TIN THƯ VIỆN**

**THÔNG TIN CHUNG**

- **Môn học:** Xử Lý Ngôn Ngữ Tự Nhiên
- **Trường:** Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh
- **Hệ đào tạo:** Thạc sĩ
- **Năm học:** 2025

**DANH SÁCH NHÓM THỰC HIỆN**

STT	Họ và Tên	Mã số Học viên
1	Phan Hoàng Khải	2531308
2	Lê Minh Nhật	2591318
3	Nguyễn Phú Thiện	2531312
4	Trần Quang Tùng	2531315

# CHƯƠNG 1: TỔNG QUAN DỰ ÁN

## 1.1. Bối cảnh và Mục tiêu

Trong khuôn khổ môn học Xử lý Ngôn ngữ Tự nhiên, dự án này tập trung vào việc xây dựng một hệ thống Chatbot dựa trên luật (Rule-Based Chatbot). Mục tiêu chính là tạo ra một công cụ trợ giúp hiệu quả, cho phép người dùng tra cứu thông tin cơ bản về thư viện một cách nhanh chóng. Đồng thời, dự án cũng là một bài tập thực tiễn giúp nhóm nghiên cứu hiểu sâu hơn về các hệ thống NLP truyền thống, phân tích những ưu và nhược điểm của chúng so với các phương pháp học máy (Machine Learning) và học sâu (Deep Learning) hiện đại.

## 1.2. Mục tiêu cụ thể

### 1.2.1. Về mặt học thuật và nghiên cứu

- Nghiên cứu và nhận diện các hạn chế cố hữu của hệ thống dựa trên luật.
- Thực hiện so sánh, đối chiếu giữa phương pháp Rule-Based và các mô hình AI hiện đại.
- Thực hành quy trình xây dựng một ứng dụng chatbot từ các thành phần cơ bản.
- Xây dựng sản phẩm minh họa (demo) về một hệ thống NLP truyền thống cho mục đích học tập.

### 1.2.2. Về chức năng của sản phẩm

- Xây dựng cơ sở tri thức gồm 12 quy tắc "từ khóa - phản hồi" (keyword-response) liên quan đến hoạt động của thư viện.
- Cung cấp thông tin chính xác về các chủ đề thường gặp như: giờ mở cửa, quy trình mượn/trả sách, và các quy định chung.
- Thiết kế giao diện trò chuyện thân thiện, trực quan và thẩm mỹ.
- Đảm bảo thiết kế đáp ứng (Responsive Design), cho phép ứng dụng hoạt động hiệu quả trên nhiều loại thiết bị (máy tính, máy tính bảng, điện thoại di động).

## 1.3. Phạm vi hỗ trợ của Chatbot

Chatbot được thiết kế để trả lời các câu hỏi liên quan đến 8 lĩnh vực chính sau:

1. **Giờ mở cửa:** Thời gian hoạt động của thư viện trong tuần và cuối tuần.
2. **Thẻ thành viên:** Thủ tục đăng ký, chi phí và các yêu cầu liên quan.
3. **Mượn và trả sách:** Quy trình, số lượng và thời hạn mượn/trả.
4. **Quy định chung:** Các nội quy về giữ im lặng, ăn uống, và an ninh.
5. **Dịch vụ hỗ trợ:** Thông tin về Wi-Fi và hệ thống máy tính công cộng.
6. **In ấn và Photocopy:** Địa điểm, chi phí và quy định sao chép tài liệu.
7. **Sử dụng không gian:** Cách thức đặt chỗ ngồi học tập và làm việc.

8. **Sự kiện:** Thông tin về các hoạt động, hội thảo do thư viện tổ chức.

## CHƯƠNG 2: CÔNG NGHỆ VÀ CẤU TRÚC

### 2.1. Công nghệ sử dụng

- **Frontend Development:** HTML5, CSS3, JavaScript (ES6+).
- **Styling:** CSS tùy chỉnh, áp dụng hiệu ứng Gradient để tăng tính thẩm mỹ cho giao diện.
- **Icons:** Thư viện Font Awesome 6.0.
- **Kiến trúc hệ thống:** Rule-Based System. Dự án không sử dụng các mô hình AI hay Machine Learning phức tạp.

### 2.2. Cấu trúc thư mục dự án

Dự án được tổ chức theo cấu trúc tệp đơn giản và rõ ràng như sau:

Rule-Based-Chatbot-NLP/

```
|— chatbot-app.html    # Tệp chính chứa giao diện ứng dụng
|— chatbot-app-styles.css # Tệp CSS định dạng giao diện
|— chatbot-app-script.js # Tệp JavaScript chứa logic xử lý
|— README.md          # Tài liệu hướng dẫn và mô tả dự án
```

### CHƯƠNG 3: CƠ SỞ TRI THỨC CỦA CHATBOT

Hệ thống chatbot hoạt động dựa trên một tập hợp gồm 12 quy tắc được định nghĩa trước. Mỗi quy tắc ánh xạ một nhóm từ khóa tới một câu trả lời tương ứng.

STT	Chủ đề	Từ khóa (Keywords)	Câu trả lời (Response)
1	<b>Giờ mở cửa</b>	giờ, mở, cửa, thời gian, hoạt động, hours, open, time, khi nào	Thư viện mở cửa từ 8:00 sáng - 10:00 tối (Thứ 2-6), 9:00 sáng - 8:00 tối (Thứ 7-CN).
2	<b>Làm thẻ</b>	làm, thẻ, đăng ký, thành viên, card, register, membership, tạo thẻ	Mang CMND/CCCD + ảnh 3x4, phí 50,000 VNĐ, xử lý trong 1-2 ngày làm việc.
3	<b>Mượn sách</b>	mượn, sách, lấy, borrow, book, lending, mượn sách	Tối đa 5 cuốn/người, thời hạn 14 ngày (sách thường), 30 ngày (giáo trình).
4	<b>Trả sách</b>	trả, sách, hoàn, return, give back, trả sách	Trả tại quầy lưu thông hoặc hộp tự động. Vui lòng kiểm tra tình trạng sách trước khi trả.
5	<b>Gia hạn sách</b>	gia hạn, kéo dài, renew, extend, thêm thời gian, gia hạn sách	Gia hạn 1 lần qua website/điện thoại, thêm 7 ngày, phí 5,000 VNĐ/cuốn.

6	<b>Phí phạt</b>	phí, phạt, tiền, fine, penalty, trả muộn, phí phạt	Trả muộn: 2,000 VNĐ/ngày/cuốn. Sách hỏng hoặc mất: đền bù gấp 2 lần giá trị bìa.
7	<b>Tìm sách</b>	tìm, sách, tìm kiếm, search, find, locate, tìm sách	Sử dụng hệ thống tra cứu trực tuyến (OPAC) trên website hoặc hỏi nhân viên tại quầy thông tin.
8	<b>Quy định</b>	quy định, nội quy, luật, rules, regulations, policy, quy định	Vui lòng không ăn uống, giữ im lặng, tắt chuông điện thoại và không mang túi lớn vào khu vực đọc sách.
9	<b>Photocopy</b>	photocopy, sao chép, in ấn, copy, print, photocopy	Dịch vụ tại tầng 1, giá 500 VNĐ/trang A4. Chỉ sao chép tối đa 30% nội dung sách.
10	<b>Wi-Fi/Máy tính</b>	wifi, internet, máy tính, computer, mạng, wifi	WiFi miễn phí (tên mạng: Library_WiFi). Có 20 máy tính tại tầng 2, sử dụng tối đa 2 giờ/phiên.
11	<b>Đặt chỗ ngồi</b>	đặt, chỗ, bàn, ghế, reserve, seat, place, đặt chỗ	Đặt chỗ qua ứng dụng hoặc website, tối đa 4 giờ/ngày. Chỗ sẽ

			bị hủy nếu đến muộn quá 15 phút.
12	<b>Sự kiện</b>	sự kiện, hoạt động, hội thảo, event, activity, seminar, sự kiện	Thư viện thường xuyên tổ chức hội thảo chuyên đề hàng tháng, triển lãm sách và các workshop kỹ năng học tập.

## CHƯƠNG 4: ĐÁNH GIÁ VÀ PHÂN TÍCH

### 4.1. Hướng dẫn sử dụng

1. Mở tệp chatbot-app.html bằng trình duyệt web.
2. Nhập câu hỏi liên quan đến thư viện vào ô trò chuyện và nhấn Gửi.
3. Ngoài ra, người dùng có thể nhấp vào các nút gợi ý nhanh (quick buttons) có sẵn.
4. Chatbot sẽ xử lý và đưa ra phản hồi dựa trên 12 quy tắc đã định nghĩa.

### 4.2. Phân tích hạn chế của phương pháp dựa trên luật

Phương pháp Rule-Based, mặc dù dễ triển khai, nhưng bộc lộ nhiều hạn chế nghiêm trọng về khả năng xử lý ngôn ngữ tự nhiên.

#### 4.2.1. Hạn chế về khả năng hiểu ngôn ngữ

- **Không hiểu ngữ cảnh:** Hệ thống không phân biệt được ý nghĩa của từ trong các ngữ cảnh khác nhau. Ví dụ, "Tôi muốn *mượn sách*" và "Cuốn sách tôi *mượn* đã cũ" đều chứa từ "mượn" nhưng mang ý định khác nhau.
- **Không xử lý từ đồng nghĩa:** Hệ thống chỉ nhận dạng các từ khóa được định nghĩa trước. Nếu người dùng sử dụng từ đồng nghĩa (ví dụ: "lấy sách" thay vì "mượn sách"), chatbot sẽ không thể hiểu.
- **Không xử lý lỗi chính tả và biến thể ngôn ngữ:** Các lỗi gõ phím, viết tắt, hoặc không dấu (ví dụ: "gio mo cua") sẽ khiến hệ thống không thể nhận dạng từ khóa.
- **Phụ thuộc hoàn toàn vào từ khóa:** Chatbot không có khả năng suy luận ý định từ các câu hỏi phức tạp không chứa từ khóa trực tiếp.

#### 4.2.2. Hạn chế về khả năng học hỏi và mở rộng

- **Không có khả năng học hỏi:** Cơ sở tri thức là tĩnh và không thể tự động cải thiện qua các lần tương tác. Mọi cập nhật đều phải được thực hiện thủ công.
- **Không hiểu ý định phức hợp:** Hệ thống không thể xử lý các câu hỏi kết hợp nhiều ý định. Ví dụ: "Tôi muốn mượn sách nhưng thẻ của tôi sắp hết hạn thì phải làm sao?" Chatbot có thể chỉ nhận diện từ khóa "mượn sách" và bỏ qua vế sau.
- **Không có bộ nhớ ngữ cảnh:** Chatbot không thể ghi nhớ các thông tin từ những câu hỏi thoại trước đó, khiến cuộc trò chuyện thiếu tự nhiên và không liền mạch.

### 4.3. Thống kê dự án

- **Số lượng quy tắc:** 12 (vượt mục tiêu ban đầu là 10).
- **Hỗ trợ ngôn ngữ:** Từ khóa hỗ trợ cả Tiếng Việt và một số thuật ngữ Tiếng Anh cơ bản.
- **Quy mô mã nguồn:** Khoảng 1,200 dòng lệnh (HTML, CSS, JS).
- **Tính tương thích:** Hoạt động tốt trên các nền tảng Mobile, Tablet, và Desktop.



## CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

### 5.1. Kết luận

Dự án đã hoàn thành các mục tiêu đề ra: xây dựng thành công một chatbot dựa trên luật với 12 quy tắc hỗ trợ tra cứu thông tin thư viện. Quan trọng hơn, sản phẩm đã minh họa một cách rõ ràng những hạn chế của phương pháp Rule-Based trong bối cảnh xử lý ngôn ngữ tự nhiên. Dự án này cung cấp một nền tảng kiến thức vững chắc, giúp nhóm nghiên cứu nhận thức rõ sự khác biệt và bước tiến hóa từ AI truyền thống sang các công nghệ AI hiện đại, qua đó tạo tiền đề cho việc tiếp cận các lĩnh vực Machine Learning và Deep Learning trong tương lai.

### 5.2. Hướng phát triển trong tương lai

Để khắc phục các nhược điểm hiện tại và nâng cao hiệu quả của chatbot, các hướng phát triển tiềm năng bao gồm:

- Tích hợp mô hình ngôn ngữ lớn (LLM):** Sử dụng các API từ OpenAI (GPT) hoặc Google (Gemini) để cải thiện khả năng hiểu và tạo sinh ngôn ngữ.
- Xây dựng mô hình Machine Learning tùy chỉnh:** Huấn luyện một mô hình phân loại ý định (Intent Classification) và nhận dạng thực thể (Entity Recognition) dựa trên dữ liệu riêng của thư viện.
- Áp dụng kỹ thuật NLP nâng cao:** Tích hợp các module xử lý tiền tố, nhận dạng từ đồng nghĩa và sửa lỗi chính tả.
- Kết nối cơ sở dữ liệu:** Sử dụng cơ sở dữ liệu để lưu trữ lịch sử hội thoại, cho phép chatbot "học hỏi" từ người dùng và cá nhân hóa trải nghiệm.
- Tích hợp API:** Kết nối với hệ thống quản lý của thư viện để cung cấp thông tin thời gian thực (ví dụ: kiểm tra tình trạng sách, gia hạn trực tiếp).

### TÀI LIỆU THAM KHẢO

- Tài liệu và giáo trình môn học Xử Lý Ngôn Ngữ Tự Nhiên.
- Các bài báo khoa học và nghiên cứu về hệ thống dựa trên luật (Rule-Based Systems).
- Tài liệu chuyên ngành về Machine Learning và Deep Learning trong NLP.
- Nghiên cứu các trường hợp ứng dụng chatbot trong lĩnh vực thư viện và dịch vụ công.