

BÁO CÁO TASK 1
“XÂY DỰNG CHATBOT DỰA TRÊN LUẬT HỖ TRỢ TRA CỨU
THÔNG TIN THƯ VIỆN”

Thành viên trong nhóm:

STT	Họ và tên	MSHV
1	Phan Hoàng Khải	2531308
2	Lê Minh Nhật	2591318
3	Nguyễn Phú Thiện	2531312
4	Trần Quang Tùng	2531315

I. TỔNG QUAN DỰ ÁN

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.

Dự án nhằm xây dựng một chatbot dựa trên luật (Rule-Based) để trả lời các câu hỏi phổ biến trong phạm vi thư viện, bao gồm: giờ mở cửa, mượn trả sách, dịch vụ wifi, và các quy định chung. Hệ thống hướng tới mục tiêu:

- Cung cấp phản hồi nhanh chóng và chính xác cho các câu hỏi thuộc phạm vi luật định sẵn.
- Là công cụ học tập hoặc demo cho khái niệm chatbot dựa trên luật, không phụ thuộc vào các mô hình AI phức tạp.
- Hỗ trợ cả tiếng Việt và tiếng Anh trong phạm vi từ khóa đã được định nghĩa.

1.1. Mục tiêu

- 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).

1.2. Phạm vi hỗ trợ của Chatbot

Chỉ các câu hỏi có thể được ánh xạ bằng từ khóa hoặc cụm từ cố định. 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.

1.3. 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.

1.4. Cấu trúc 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

1.5. 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	Giờ mở cửa	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	Làm thẻ	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	Mượn sách	mượn, sách, lấy, borrow, book, lending, mượn sách	Tối đa 5 cuốn/người, thời hạn 14 ngày (sách thường), 30 ngày (giáo trình).
4	Trả sách	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	Gia hạn sách	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	Phí phạt	phí, phạt, tiền, fine, penalty, trả mượn, phí phạt	Trả mượn: 2,000 VNĐ/ngày/cuốn. Sách hỏng hoặc mất: đền bù gấp 2 lần giá trị bìa.

7	Tìm sách	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	Quy định	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	Photocopy	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	Wi-Fi/Máy tính	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	Đặt chỗ ngồi	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	Sự kiện	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.

1.6. Hướng dẫn sử dụng

Bước 1: Mở tệp chatbot-app.html bằng trình duyệt web.

Bước 2: Nhập câu hỏi liên quan đến thư viện vào ô trò chuyện và nhấn Gửi.

Bước 3: Ngoài ra, người dùng có thể nhấp vào các nút gợi ý nhanh (quick buttons) có sẵn.

Bước 4: Chatbot sẽ xử lý và đưa ra phản hồi dựa trên 12 quy tắc đã định nghĩa.

II. THẢO LUẬN

2.1. Hạn chế

2.2.1. Hạn chế chung

- Thiếu linh hoạt, dễ lỗi khi thay đổi: Chỉ phản hồi nếu câu hỏi khớp từ khóa được định nghĩa. Và cần người dùng viết hơi khác đi là có thể không nhận ra.

- Không hiểu nghĩa ngoài từ khóa: Không phân tích ngữ cảnh hay mối quan hệ giữa các từ. Ví dụ: “Tôi muốn mượn sách nhưng thẻ hết hạn” và “thẻ hết hạn, có mượn được không?” có thể không khớp luật nào.

- Không học được (không thích nghi): Nếu có câu hỏi mới, chatbot không tự cải thiện, chỉ có thể cập nhật thủ công.

- Xử lý kém với liên tiếp câu hỏi cùng 1 nội dung hoặc ngữ cảnh: Không nhớ hội thoại trước đó, nên câu tiếp nối (“thế cuối tuần thì sao?”) sẽ không giải quyết được. Một từ có nhiều nghĩa (đa nghĩa) dễ gây nhầm lẫn.

- Khó mở rộng: Khi số lượng luật nhiều lên, việc quản lý và duy trì sẽ phức tạp. Có thể xung đột giữa các luật trùng từ khóa.

- Trải nghiệm người dùng kém: Nếu chatbot thường xuyên trả lời “Tôi không hiểu” sẽ gây khó chịu. Trả lời khô khan, thiếu tự nhiên.

2.2.2. Phân tích hạn chế

** Hạn chế tức thời:*

Hạn chế	Phân tích	Ví dụ minh họa
Thiếu linh hoạt	Chatbot chỉ phản hồi đúng khi người dùng nhập chính xác từ khóa có trong luật. Chỉ cần thay đổi cách diễn đạt, bot sẽ không nhận diện.	Luật định nghĩa: “mượn sách” → bot trả lời. Người dùng gõ: “lấy sách” → bot không phản hồi.
Không xử lý được lỗi chính tả	Sai chính tả hoặc gõ không dấu sẽ khiến hệ thống không khớp được từ khóa.	Người dùng gõ “muon sach” thay vì “mượn sách” → bot không hiểu.

Phản hồi hạn chế	Nếu không có luật khớp, bot thường trả về “Tôi không hiểu”, gây trải nghiệm khô khan, lặp lại.	Người dùng hỏi: “Có thể mượn sách tham khảo không?” (không có luật) → bot trả “Tôi không hiểu”.
-------------------------	--	---

*** Hạn chế nghiêm trọng:**

Hạn chế	Phân tích	Ví dụ minh họa
Không hiểu ngữ cảnh hội thoại	Bot không duy trì trạng thái; không theo dõi các câu hỏi trước đó, dẫn đến mất mạch hội thoại.	Người dùng: “Thư viện mở cửa đến mấy giờ?” → bot trả lời: “Đến 21h.”. Người dùng tiếp: “Thế cuối tuần thì sao?” → bot không hiểu câu hỏi nối tiếp.
Khó mở rộng và bảo trì	Khi số lượng luật nhiều, dễ trùng lặp, xung đột, và tốn công quản lý.	Có nhiều luật liên quan đến “mượn sách” ở nhiều tình huống khác nhau → việc sửa hoặc bổ sung rất phức tạp.
Không có khả năng tự học	Hệ thống không tự cải thiện sau khi gặp các câu hỏi mới, bắt buộc phải cập nhật thủ công.	Người dùng hỏi liên tục về “gia hạn thẻ online” → bot không học được mà phải thêm luật thủ công.
Giới hạn trải nghiệm người dùng	Bot thường xuyên không hiểu, dễ khiến người dùng mất kiên nhẫn, giảm sự tin tưởng.	Sau 3–4 lần hỏi khác nhau mà bot đều trả lời “không hiểu”, người dùng có xu hướng bỏ cuộc.

2.3. Xử lý “từ đồng nghĩa”, “lỗi chính tả” và “câu hỏi chưa theo quy tắt”

Tình huống	Trường hợp	Ví dụ	Hệ thống hiện tại	Giải pháp đề xuất
Từ đồng nghĩa	1. Khác từ nhưng cùng nghĩa trong tiếng Việt	“mượn sách” (được hỗ trợ) vs “lấy sách” (không được hỗ trợ)	Không khớp → bot trả “không hiểu”.	Bổ sung danh sách từ đồng nghĩa (mượn = lấy = borrow).
	2. Dùng tiếng Anh hoặc ngôn ngữ khác	“borrow book” vs “loan book”	Không khớp nếu chỉ định nghĩa nghĩa “borrow”.	Thêm mapping đa ngôn ngữ, hoặc dùng từ điển synonym.
	3. Dùng từ địa phương/thông tục	“xin sách” thay cho “mượn sách”	Không khớp từ khóa.	Bổ sung từ điển mở rộng theo ngữ cảnh/địa phương.
Lỗi chính tả	1. Sai chính tả đơn giản	“muon” thay vì “mượn”	Không khớp từ khóa.	Fuzzy matching (Levenshtein ≤ 2 ký tự).
	2. Lỗi gõ do bộ gõ/telex	“muon sach” thay vì “mượn sách”	Không nhận diện.	Chuẩn hóa văn bản trước khi khớp (bỏ dấu, loại bỏ stopword).
	3. Lỗi chính tả phức tạp	“moun” thay cho “mượn”	Bỏ qua.	Dùng spell-checker hoặc model gợi ý từ gần đúng.

Câu hỏi chưa theo quy tắt	1. Câu hỏi mở rộng ngoài phạm vi	“Có thể gia hạn thẻ online không?”	Không có luật → bot trả “không hiểu”.	Fallback: hiển thị danh sách dịch vụ có sẵn.
	2. Câu hỏi kết hợp nhiều ý	“Mượn sách và đăng ký wifi thẻ nào?”	Bot chỉ bắt được 1 luật hoặc không bắt được.	Phân tách câu, xử lý từng ý rồi gộp trả lời.
	3. Câu hỏi hoàn toàn mới	“Có quán cà phê trong thư viện không?”	Không có trong luật.	Bot xin lỗi + hướng người dùng liên hệ nhân viên, hoặc chuyển tiếp đến kênh hỗ trợ.

III. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

3.1. 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.

Có thể thấy, chatbot dựa trên luật có cấu trúc đơn giản nhưng lại bộc lộ nhiều hạn chế cơ bản: không xử lý được : từ đồng nghĩa, biến thể ngôn ngữ, lỗi chính tả, câu hỏi ngoài phạm vi. Ngoài ra còn các hạn chế nghiêm trọng khác như là: thiếu khả năng hiểu ngữ cảnh, thiếu khả năng ghi nhớ hội thoại, thiếu khả năng học hỏi và mở rộng; bất lực trước các yêu cầu kết hợp.

Những hạn chế này không chỉ cản trở hiệu quả của ứng dụng chatbot mà còn làm giảm đáng kể trải nghiệm của người dùng.

Vì vậy, phương pháp này khó đáp ứng được yêu cầu thực tế trong các hệ thống giao tiếp hiện đại, và thường chỉ phù hợp cho các ứng dụng thử nghiệm quy mô nhỏ hoặc trong phạm vi nghiệp vụ hẹp.

3.2. Hướng phát triển

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).