**CUỘC THI NGHIÊN CỨU KHKT DÀNH CHO HS TRUNG HỌC**

**NĂM HỌC 2024-2025**

**BÁO CÁO TÓM TẮT DỰ ÁN:**

**Phát triển mô hình trợ lý Robot thủ thư ứng dụng**

**mô hình ngôn ngữ lớn và thị giác máy tính**

**Lĩnh vực dự thi: Robot và máy thông minh**

**Tháng 12/2024**

**MỤC LỤC**

**I. Lý do chọn đề tài 1**

**II. Mục tiêu đề tài 4**

**III. Thiết kế và phương pháp nghiên cứu**

**1 Tổng quan 4**

**1.1 Tính năng hỏi đáp và gợi ý bạn đọc**

**1.2.1 Tổng quan về hệ thống gợi ý 4**

**1.2.2 Tổng quan về chatbot hỏi đáp dựa trên LLM và RAG 5**

**2 Phần mềm**

**2.1 Nguyên lý hoạt động chung**

**2.2 Phát triển hệ thống hỏi đáp và gợi ý sách 5**

**2.2.1 Thu thập dữ liệu sách tiếng Việt và tiếng Anh**

**2.2.2 Mô hình mã hóa dữ liệu văn bản**

**2.2.3 Thử nghiệm mô hình ngôn ngữ lớn với cơ sở dữ liệu đã có**

**2.2.4 Đánh giá hệ thống và hiệu chỉnh**

**2.3 Tính năng truy xuất sách bằng thị giác máy**

**2.4 Tính năng giao tiếp bằng giọng nói 5**

**3 Phần cứng 6**

**3.1 Các thành phần chính 6**

**3.2 Nguyên lý hoạt động chung 8**

**6.3. Thiết kế, in 3D và gia công vỏ robot 9**

**6.3.1. Nguyên mẫu 1 10**

**6.3.2. Nguyên mẫu 2 12**

**7. Thử nghiệm và đánh giá hệ thống 14**

**IV. Kết luận 12**

**1. Đánh giá kết quả 15**

**1.1 Kết quả 15**

**1.2 Hạn chế 15**

**2. Hướng phát triển 15**

**2.1. Nâng cấp hệ thống camera 15**

**2.2. Phát triển tính năng 16**

**16**

**Tháng 12/2024**

**I. Lý do chọn đề tài**

Thư viện từ lâu đã được xem là điểm đến kho tàng tri thức, nơi lưu giữ và cung cấp các nguồn thông tin phong phú, có tổ chức, đáp ứng nhu cầu học tập và nghiên cứu của mọi đối tượng độc giả. Tuy nhiên, để hoàn thành sứ mệnh cao cả đó, các thư viện phải trải qua một loạt quy trình kỹ thuật phức tạp, từ việc biên mục, phân loại tài liệu theo các tiêu chuẩn quốc tế cho đến tổ chức không gian và hướng dẫn người dùng tiếp cận. Đối với người sử dụng, việc khai thác tài nguyên thư viện không phải lúc nào cũng dễ dàng. Trong các thư viện vật lý, người đọc thường phải thực hiện nhiều bước để tiếp cận tài liệu mong muốn. Đầu tiên, họ cần sử dụng hệ thống thông tin quản lý thư viện để xác định tài nguyên (sách mềm hoăc sách cứng, bản ghi,...) có sẵn và ghi chú thông tin như số đăng ký hoặc vị trí kệ sách. Sau đó, họ phải di chuyển trong không gian thư viện theo các biển chỉ dẫn, tìm đến khu vực được phân loại theo chủ đề hoặc chuyên ngành, và tiếp tục tìm kiếm trên các kệ sách. Quy trình này, tuy khoa học và có tổ chức, nhưng đôi khi lại trở nên phức tạp và gây khó khăn, đặc biệt với những người không quen thuộc hoặc lần đầu sử dụng thư viện.

Hơn nữa, đối tượng học sinh – một trong những nhóm độc giả chính của thư viện trường học – thường không có đủ thời gian hoặc kỹ năng để tra cứu và tìm kiếm tài liệu theo cách truyền thống. Điều này không chỉ làm giảm hiệu quả học tập mà còn khiến nhiều học sinh mất dần sự hứng thú với việc tự học qua tài liệu sách.Từ thực trạng này, nhu cầu cấp thiết về một giải pháp hỗ trợ hiện đại hóa thư viện, giúp đơn giản hóa quy trình tìm kiếm và quản lý tài nguyên, trở nên rõ ràng hơn bao giờ hết. Một hệ thống thông minh như robot thủ thư với khả năng giao tiếp tự nhiên, tìm kiếm thông tin chính xác và hỗ trợ cá nhân hóa có thể là chìa khóa để giải quyết vấn đề này.

Việc ứng dụng robot trong thư viện đã được triển khai tại nhiều nơi trên thế giới, trong đó bao gồm cả Việt Nam. Điển hình là các robot như robot Tu Xiaoling, đã được đưa vào sử dụng tại thư viện Thành phố Thượng Hải, Trung Quốc vào năm 2022. robot này được trang bị tính năng tư vấn các chủ đề thông thường cho độc giả, có thể trả lời câu hỏi đơn giản thông qua giao tiếp bằng ngôn ngữ tiếng Trung hoặc tiếng Anh.

**

Hình 1: Robot trợ lý thư viện Tu Xiaoling tại Thượng Hải

Ưu điểm:

+) Thiết kế gần gũi, tăng cảm giác thích thú với người dùng.

+) Tính năng giao tiếp thông minh, có thể trả lời câu hỏi về sách.

Nhược điểm:

+) Hạn chế về ngôn ngữ giao tiếp.

+) Khả năng hiểu các yêu cầu người dùng trong ngữ cảnh phức tạp kém do không áp dụng các mô hình ngôn ngữ lớn hiện đại như chatGPT, Llama hay Claude.

+) Không thể gợi ý sách bạn đọc dựa vào nhu cầu và sở thích.

Hay nếu bạn cần tìm cuốn sách trong một thư viện lớn có thể là một nhiệm vụ khó khăn, nhưng giờ đây công việc này trở nên dễ dàng hơn tại một thư viện ở tỉnh Chiết Giang, Trung Quốc nhờ cánh tay robot tự động tìm và lấy/trả sách cho người dùng.



Hình 2: Cánh tay robot phục vụ lấy/trả sách

trong thư viện tỉnh Triết Giang, Trung Quốc

Ưu điểm:

+)Giúp tự động hóa và rút ngắn phần lớn quy trình mượn, trả sách của người dùng

Nhược điểm:

+) Chi phí đầu tư lớn.

+) Không thể gợi ý sách.

Tại Việt Nam cũng đã có một số cơ sở và tổ chức ứng dụng Robot phục vụ trong thư viện. Mới nhất gần đây là Mô Hình Robot Thư Viện của nhóm nghiên cứu đại học Trường Đại học Sư phạm Kỹ thuật Thành phố Hồ Chí Minh kết hợp với Công ty Cổ phần Optima Robotics, Việt Nam[1]. Đây là một robot di động đa năng hoạt động phục vụ trong thư viện bao gồm các tính năng: Nhận diện khuôn mặt người dùng thay thế cho thẻ thư viện, tính năng giao tiếp đơn giản, tính năng tìm kiếm sách qua mã vạch cơ bản, tính năng tự di chuyển tìm sách.



Hình 3: Mô hình Robot phục vụ thư viện tại Trường đại học

Sư phạm Kỹ Thuật Thành phố Hồ Chí Minh

Ưu điểm:

+)Giúp tự động hóa và rút ngắn phần lớn quy trình mượn, trả sách của người dùng.

+) Hỗ trợ giao tiếp tiếng Việt.

Nhược điểm:

+) Thiết kế không gần gũi với người dùng.

+) Không thể gợi ý sách dựa trên nhu cầu và sở thích người đọc.

Qua nghiên cứu các giải pháp đã có như trên và nhiều giải pháp khác nhau. Nhóm nhận thấy việc thiếu vắng các giải pháp tập trung vào trải nghiệm người dùng khiến thư viện vẫn mang tính truyền thống và thụ động trong mắt độc giả trẻ, đặc biệt là học sinh. Tuy các robot trên có thể phục vụ việc tìm sách và tra cứu, nhưng người dùng phải tự mình đánh giá mức độ phù hợp của tài liệu với nhu cầu cá nhân hay muốn hỏi đáp thông tin về nội dung của sách còn hạn chế. Những khó khăn này không chỉ làm giảm hiệu quả sử dụng thư viện mà còn khiến nhiều người, đặc biệt là học sinh, mất dần sự hứng thú với việc tự học.

**II. Mục tiêu đề tài**

Đề tài robot thủ thư ứng dụng mô hình ngôn ngữ lớn và thị giác máy tính ra đời với mục tiêu khắc phục những hạn chế hiện tại, nhằm mang đến một giải pháp đột phá giúp nâng cao trải nghiệm người dùng. Robot không chỉ hỗ trợ tra cứu và kiểm kê sách như các hệ thống truyền thống, mà còn hoạt động như một người trợ lý thông minh, sẵn sàng gợi ý tài liệu dựa trên sở thích cá nhân hoặc hỏi đáp về nội dung mà người dùng quan tâm bằng cách giao tiếp giọng nói tiếng Việt hoặc thông qua giao diện chat (nhắn tin) trên màn hình. Ngoài ra, người dùng có thể dễ dàng check-in lấy/trả sách nhanh chóng và thông minh qua camera nhận diện sách tự động đặt trên robot. Đặc biệt với tiêu chí đơn giản, dễ sử dụng, có thể triển khai hệ thống trên các thiết bị máy tính nhúng cấu hình thấp nhỏ gọn tiết kiệm năng lượng và xử lý nhanh phù hợp để ứng dụng thực tế trong các thư viện.

**III. Thiết kế và phương pháp nghiên cứu**

**1. Tổng quan**

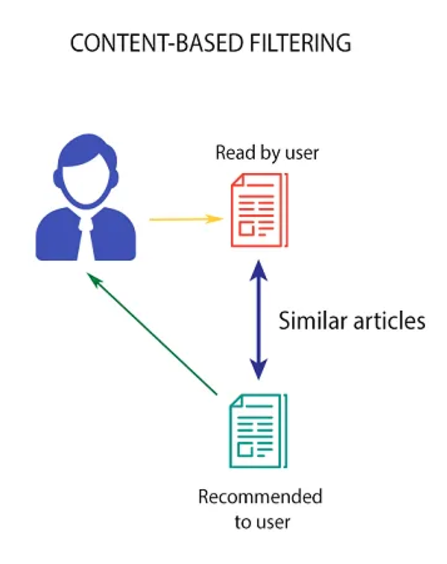
**1.1. Tổng quan về hệ thống gợi ý**

Ngày nay, người dùng phải đối mặt với việc lựa chọn từ hàng triệu sản phẩm, nội dung và dịch vụ trên các nền tảng trực tuyến. Chính vì vậy, hệ thống gợi ý, vốn cung cấp các đề xuất phù hợp và cá nhân hóa nhất cho người dùng, ngày càng trở nên quan trọng. Mục tiêu tổng quát của hệ thống gợi ý là lọc ra từ tập hợp khổng lồ các nội dung, sản phẩm sách mềm, tài liệu số, bản ghi, v.v., những thứ phù hợp nhất với sở thích và nhu cầu của từng cá nhân. Điều này không chỉ giúp người dùng tiết kiệm thời gian mà còn tối ưu hóa trải nghiệm của họ trong thư viện.

**1.2. Các mô hình hệ thống gợi ý truyền thống**

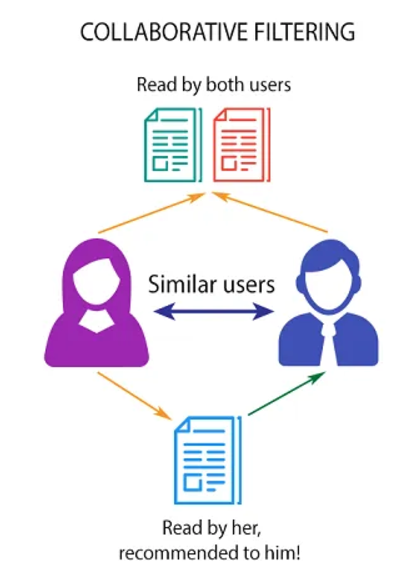
**1.2.1. Hệ thống gợi ý Lọc nội dung (Content-Based Filtering)**

Hệ thống gợi ý dựa trên nội dung (Content-Based Recommender System) hoạt động bằng cách dự đoán các đặc điểm hoặc hành vi của người dùng dựa trên các đặc tính của những mục mà người dùng đã phản hồi tích cực. Nói cách khác, hệ thống sẽ phân tích đặc điểm của sản phẩm hoặc nội dung mà người dùng quan tâm, sau đó tìm kiếm các mục tương tự để đưa ra gợi ý.



Hình 3: Mô tả nguyên lý hệ thống gợi ý dựa trên Lọc nội dung

**1.2.2. Lọc cộng tác (Collaborative Filtering)**

****

Hình 4: Mô tả nguyên lý hệ thống gợi ý dựa trên lọc cộng tác

Phương pháp lọc cộng tác không yêu cầu các đặc điểm của sản phẩm hoặc nội dung phải được cung cấp trước. Thay vào đó, hệ thống tự động tạo các vector đặc trưng (embedding) cho cả người dùng và sản phẩm, đồng thời ánh xạ chúng vào cùng một không gian embedding. Hệ thống này dựa trên phản hồi từ các người dùng khác khi đưa ra gợi ý cho một người dùng cụ thể. Nó phân tích các mục mà một người dùng thích, sau đó xem xét các mục mà những người dùng có hành vi và sở thích tương tự cũng thích để đưa ra các gợi ý phù hợp. Quá trình này lặp lại nhiều lần để giảm sai số giữa dự đoán và dữ liệu thực tế, từ đó cải thiện độ chính xác.

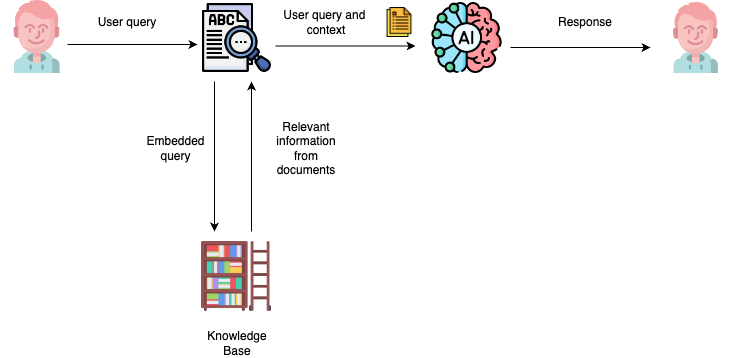
Ví dụ minh họa: Khi người dùng đánh giá sách, có những yếu tố không được chỉ rõ như cảm nhận về nội dung, kịch bản, hoặc thể loại. Những yếu tố này cũng tồn tại tương tự ở các loại sách. Khi hệ thống tìm thấy mối quan hệ giữa các yếu tố tiềm ẩn này, nó có thể dự đoán mức độ yêu thích của người dùng với các cuốn sách khác.

**2. Phát triển hệ thống gợi ý sách sử dụng mô hình ngôn ngữ lớn**

Trong bối cảnh khám phá sách không ngừng phát triển, các hệ thống gợi ý truyền thống thường tỏ ra hạn chế. Các mô hình ngôn ngữ lớn (LLMs) mang đến một cách tiếp cận đầy triển vọng. Nhờ khả năng xử lý khối lượng dữ liệu văn bản khổng lồ, LLMs có thể phân tích sâu sắc các thể loại, phong cách viết và sở thích của độc giả. Chính sự sâu sắc mới mẻ này mở ra tiềm năng cách mạng hóa việc gợi ý sách, không chỉ dẫn dắt độc giả đến những mô-típ quen thuộc mà còn thực sự được cá nhân hóa. Mở rộng cơ hội khám phá những tác phẩm độc đáo, vượt xa mong đợi của họ. Điều này đặt nền tảng cho một thế hệ công cụ gợi ý mới, tối ưu hóa trải nghiệm của độc giả và góp phần thúc đẩy việc tiếp cận tri thức toàn diện hơn.

**2.1. Nguyên lý hoạt động chung**

Một trong những tiến bộ đáng chú ý trong các hệ thống gợi ý sách sử dụng mô hình ngôn ngữ lớn (LLM) là sự tích hợp của Retrieval-Augmented Generation (RAG). RAG hoạt động như một công cụ truy xuất thông tin tiên tiến dành cho LLM. Bằng cách tìm kiếm hiệu quả trong các tập dữ liệu khổng lồ về thông tin sách, RAG có thể xác định các tựa sách có nội dung và yếu tố phong cách tương tự. Dữ liệu truy xuất này giúp LLM vượt ra khỏi các gợi ý dựa trên sự tương đồng đơn giản. Nhờ có RAG, LLM không chỉ hiểu rõ các chủ đề cơ bản và cách tiếp cận câu chuyện mà còn có thể kết nối chúng với sở thích cá nhân của người dùng.

****

Hình 5: Ảnh mô tả nguyên lý hoạt động hệ thống gợi ý kết hợp LLMS và RAG

Trong đó:

**User query:** Thay vì một câu hỏi mở, người dùng có thể đưa ra các thông tin như: thể loại sách yêu thích, tác giả, chủ đề muốn tìm hiểu, hoặc thậm chí chỉ là một cảm xúc muốn được truyền tải qua sách.

**Embedded query:** Câu hỏi của người dùng được chuyển đổi thành một vector số học, đại diện cho ý nghĩa của câu hỏi đó trong không gian vector.

**Knowledge base:** Thay vì một cơ sở dữ liệu văn bản đơn thuần, hệ thống sẽ có một thư viện vector sách khổng lồ, thông tin mỗi cuốn sách được biểu diễn dưới dạng một vector.

**Relevant information:** Hệ thống sẽ tìm kiếm những cuốn sách có vector gần nhất với vector của câu hỏi người dùng, tức là những cuốn sách có nội dung phù hợp nhất với yêu cầu.

**AI:** Mô hình LLM sẽ sử dụng thông tin từ những cuốn sách được tìm thấy để tạo sinh ra văn bản với nội dung là một gợi ý đọc sách cụ thể, có thể bao gồm: tên sách, tác giả, tóm tắt nội dung, lý do tại sao cuốn sách này phù hợp với người dùng.

**Response:** Cuối cùng, hệ thống sẽ trả về cho người dùng một danh sách các gợi ý đọc sách, cùng với những thông tin bổ sung để người dùng có thể đưa ra quyết định.

**2.2. Phát triển hệ thống hỏi đáp và gợi ý sách**

**2.2.1. Thu thập dữ liệu sách tiếng Việt và tiếng Anh**

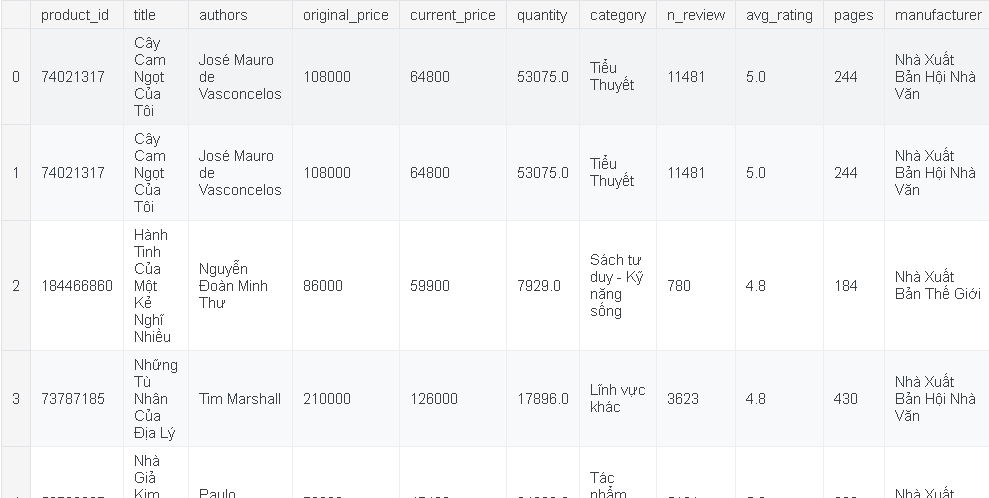
Thu thập dữ liệu về sách là bước quan trọng để xây dựng hệ thống gợi ý cho robot thư viện. Với vai trò hỗ trợ người dùng tìm kiếm sách hiệu quả, robot cần tích hợp dữ liệu từ hai nguồn chính: kho sách thực tế của thư viện và các cơ sở dữ liệu sách trực tuyến. Dữ liệu nguồn sách trực tuyến cung cấp các bản tóm tắt, đánh giá, tác giả, năm xuất bản và thông tin chi tiết bổ sung, còn từ thư viện sẽ bao gồm thêm thông tin về vị trí, danh mục, và mã vạch của từng sách. Sự kết hợp này đảm bảo rằng hệ thống không chỉ quản lý tốt tài nguyên tri thức sẵn có nội bộ mà còn mở rộng khả năng gợi ý với các đầu sách phong phú trên mạng. Nguồn sách, bản ghi chỉ xuất bản nội bộ sẽ được thêm thủ công bằng cách nhập liệu vào phần mềm của hệ thống. Với nguồn sách trên mạng, ta có thể cào dữ liệu trên các trang bán sách thương mại điện tử như shopee, tiki hoặc sử dụng bộ dữ liệu sách sẵn có mã nguồn mở trên mạng từ các trang web chuyên về dữ liệu khoa học như kaggle, hugging face hub, . . .



Hình 6: Cào dữ liệu sách trên mạng internet

Nhóm lựa chọn trang thương mại điện tử tập trung về sách Tiki để thu thập dữ liệu về sách. Từ trang web này, sử dụng thư viện Python,Requests,BeautifulSoup,Pandas. chúng ta sẽ lấy thông tin sau: “Tên sách, tác giả, giá thành, thể loại, nhà xuất bản, mô tả sách”.

Sau khi thu thập thông tin, chúng ta sẽ lưu trữ thông tin đó trong Pandas Dataframe và chuyển đổi thành tệp .CSV



Hình 7: Ví dụ một vài dữ liệu mẫu sách sau khi thu thập

**2.2.2. Mô hình mã hóa dữ liệu văn bản**

Sau khi thu thập được dữ liệu văn bản, bước tiếp theo là chuyển đổi chúng thành các vector mã hóa (embedded vectors) để máy tính có thể hiểu và xử lý. Quá trình này sử dụng các mô hình nhúng (embedding models) để biểu diễn văn bản dưới dạng các vector số trong không gian đa chiều. Các vector này không chỉ lưu trữ thông tin cơ bản mà còn phản ánh các mối quan hệ ngữ nghĩa, chẳng hạn như sự tương đồng giữa nội dung của các cuốn sách hoặc giữa sở thích của người dùng và đặc điểm của sách. Đây là nền tảng để tính toán mức độ tương đồng, xếp hạng, và gợi ý nội dung phù hợp.

**2.2.3. Ứng dụng mô hình ngôn ngữ lớn với cơ sở dữ liệu đã có**

**2.2.3.1. Nền tảng Langchain**

****

Hình 8: Khung mã nguồn mở LangChain

LangChain là một khung mã nguồn mở để xây dựng các ứng dụng dựa trên các mô hình ngôn ngữ lớn (LLM). LLM là các mô hình học sâu lớn được đào tạo trước trên khối lượng lớn dữ liệu có thể tạo ra câu trả lời cho các câu hỏi của người dùng, ví dụ như trả lời câu hỏi hoặc tạo hình ảnh từ lời nhắc dựa trên văn bản. LangChain cung cấp các công cụ và yếu tố trừu tượng để cải thiện khả năng tùy chỉnh, độ chính xác và mức độ liên quan của thông tin do các mô hình tạo ra. Ví dụ: nhà phát triển có thể sử dụng các thành phần LangChain để xây dựng chuỗi nhắc mới hoặc tùy chỉnh các mẫu hiện có. LangChain cũng bao gồm các thành phần cho phép LLM truy cập các tập dữ liệu mới mà không cần đào tạo lại.

**2.2.3.2 Mô hình ngôn ngữ lớn**

Hiện nay, có rất nhiều công ty và nền tảng cung cấp API cho các mô hình ngôn ngữ lớn (LLM) đã được đào tạo sẵn, hỗ trợ mạnh mẽ trong việc tạo nhúng (embedding) và tạo văn bản (generation). Các API này giúp tiết kiệm thời gian và công sức trong việc xây dựng mô hình từ đầu, đồng thời cung cấp các công cụ tối ưu và hiệu quả cao.

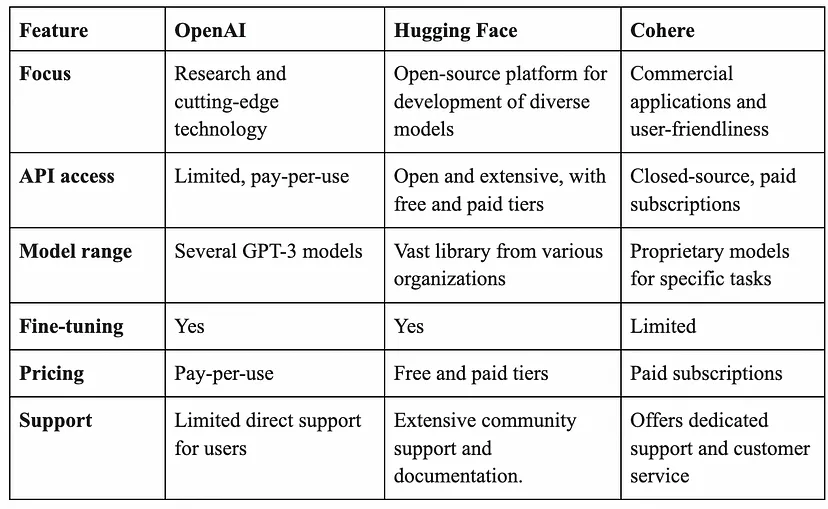
Một số công ty và nền tảng phổ biến bao gồm:

OpenAI: Cung cấp các mô hình GPT mạnh mẽ như GPT-4 và GPT-3.5, hỗ trợ cả việc tạo văn bản lẫn nhúng dữ liệu với API dễ sử dụng.

HuggingFace: Một nền tảng mã nguồn mở hàng đầu, cung cấp hàng nghìn mô hình pretrained như BERT, RoBERTa, và DistilBERT, phù hợp với cả nhúng và các tác vụ NLP khác.

Cohere: Tập trung vào các dịch vụ NLP và tạo embedding chất lượng cao, với API thân thiện và khả năng tùy chỉnh linh hoạt.

Những nền tảng này không chỉ giúp triển khai nhanh các giải pháp mà còn hỗ trợ tài liệu, ví dụ minh họa, và tích hợp dễ dàng với các ngôn ngữ lập trình phổ biến, mở ra cơ hội tối ưu hóa hệ thống gợi ý một cách nhanh chóng và hiệu quả.

****

Hình 9: Các nền tảng LLM phổ biến

Trong quá trình thử nghiệm, OpenAI đã được sử dụng với các tín dụng miễn phí ban đầu. Tuy nhiên, do kích thước lớn của tập dữ liệu, hệ thống đã gặp phải giới hạn tốc độ (rate limit) mỗi khi thực hiện các thao tác phức tạp như tạo nhúng (embedding) hoặc xử lý khối lượng lớn thông tin. Do đó, Hugging Face đã được chọn làm nền tảng LLM chính. Hugging Face không chỉ cung cấp các mô hình mạnh mẽ và đa dạng, mà còn mang lại sự linh hoạt hơn trong việc xử lý dữ liệu quy mô lớn mà không bị hạn chế bởi giới hạn tốc độ. Nền tảng này phù hợp với yêu cầu về cả hiệu năng và khả năng tùy chỉnh, giúp triển khai hệ thống hiệu quả hơn.

**2.2.3.3 Tổng quan về cơ sở dữ liệu vector**

Cơ sở dữ liệu vector (Vector Database) là một loại cơ sở dữ liệu được thiết kế đặc biệt để lưu trữ và tìm kiếm các vector nhúng (embedded vectors). Vector đại diện cho các dữ liệu phức tạp như văn bản, hình ảnh, hoặc âm thanh trong không gian đa chiều. Điều này rất hữu ích trong các ứng dụng như tìm kiếm ngữ nghĩa, gợi ý cá nhân hóa, và xử lý ngôn ngữ tự nhiên. Cơ sở dữ liệu vector giúp thực hiện các phép toán so sánh, tìm kiếm, hoặc truy vấn dựa trên độ tương đồng (similarity) giữa các vector, sử dụng các kỹ thuật như Cosine Similarity, Euclidean Distance, hoặc các thuật toán tối ưu khác. Điều này cải thiện hiệu quả và tốc độ xử lý khi làm việc với khối lượng lớn dữ liệu nhúng. Với việc sử dụng LangChain cung cấp các trình kết nối (connectors) tích hợp với nhiều cơ sở dữ liệu vector như: Chroma, FAISS (Facebook AI Similarity Search), Pinecone. Trong dự án này, nhóm đã sử dụng so sánh hai công cụ phổ biến: Chroma và FAISS:

Chroma: Một cơ sở dữ liệu vector nguồn mở, dễ triển khai và phù hợp cho các dự án quy mô vừa. Nó cung cấp giao diện thân thiện và tích hợp tốt với LangChain.

FAISS: Một thư viện mạnh mẽ được phát triển bởi Facebook AI, tối ưu hóa cho tìm kiếm vector với hiệu suất cao trên các bộ dữ liệu lớn, sử dụng tốt cho các tác vụ đòi hỏi độ chính xác cao.

Việc so sánh giúp xác định công cụ phù hợp nhất, cân bằng giữa yêu cầu về hiệu năng, độ phức tạp và khả năng triển khai trong hệ thống.

**2.3 Tính năng giao tiếp bằng giọng nói**

Để nâng cao khả năng tương tác của robot thủ thư, hệ thống được tích hợp tính năng giao tiếp bằng giọng nói, cho phép người dùng giao tiếp tự nhiên với robot mà không cần sử dụng bàn phím hay màn hình cảm ứng. Tính năng này được triển khai thông qua:

Thư viện API Google TTS (Text-to-Speech) Google TTS chuyển đổi văn bản thành giọng nói tự nhiên, giúp robot phản hồi lại người dùng bằng âm thanh. Phù hợp để thông báo vị trí sách, hướng dẫn mượn/trả sách, hoặc cung cấp thông tin gợi ý.

Thư viện API Google STT (Speech-to-Text): Google STT chuyển đổi giọng nói của người dùng thành văn bản, giúp robot hiểu và xử lý câu hỏi hoặc yêu cầu.

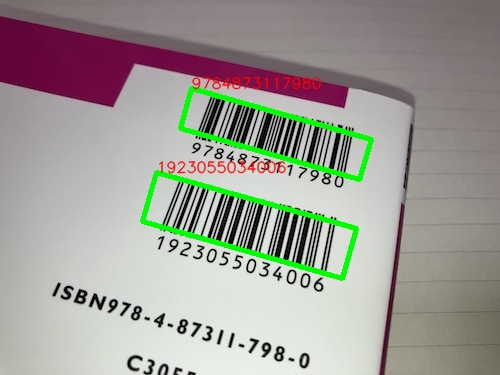
Ưu điểm

Tăng tính tương tác, phù hợp với người dùng chưa quen sử dụng giao diện số.

Hỗ trợ người dùng khuyết tật hoặc không thể sử dụng màn hình cảm ứng.

Mang lại trải nghiệm hiện đại, phù hợp với xu hướng công nghệ tiên tiến.

**2.4. Tính năng nhận diện mã vạch sử dụng thị giác máy**

Để hỗ trợ quy trình quản lý mượn và trả sách một cách nhanh chóng, robot thư viện được trang bị tính năng nhận diện mã vạch sách dựa trên công nghệ thị giác máy. OpenCV, một thư viện xử lý ảnh mạnh mẽ và phổ biến, được sử dụng để phát triển tính năng này.

Hình 10: Ví dụ về sử dụng thị giác

máy nhận diện mã vạch

**2.4.1. Nguyên lý hoạt động chung**

Bước 1: Thu thập hình ảnh từ camera: Khi người dùng đưa sách vào tầm nhìn của camera, hệ thống sẽ ghi lại hình ảnh

Bước 2: Tiền xử lý hình ảnh: OpenCV thực hiện các bước như chuyển đổi ảnh sang thang độ xám, lọc nhiễu, và tăng cường độ tương phản để tối ưu hóa khả năng nhận diện.

Bước 3: Phát hiện mã vạch:

Hệ thống sử dụng các thuật toán dò tìm cạnh và phân tích hình dạng để xác định vùng chứa mã vạch.

Khi phát hiện vùng mã vạch, OpenCV sẽ tách riêng vùng này để phân tích.

Bước 4: Giải mã mã vạch:

OpenCV tích hợp các thư viện hỗ trợ đọc mã vạch như ZBar hoặc các công cụ giải mã mã vạch khác.

Dữ liệu mã vạch được trích xuất và chuyển đổi thành thông tin sách tương ứng (ID sách, tên sách, v.v.).

Bước 5: Kết nối với cơ sở dữ liệu:

Mã vạch đã giải mã được gửi đến hệ thống quản lý thư viện để tra cứu và cập nhật trạng thái sách (mượn, trả, hoặc kiểm tra thông tin).

Ưu điểm của giải pháp

Tự động hóa: Giảm thiểu sự can thiệp thủ công, tăng tốc độ xử lý mượn/trả sách.

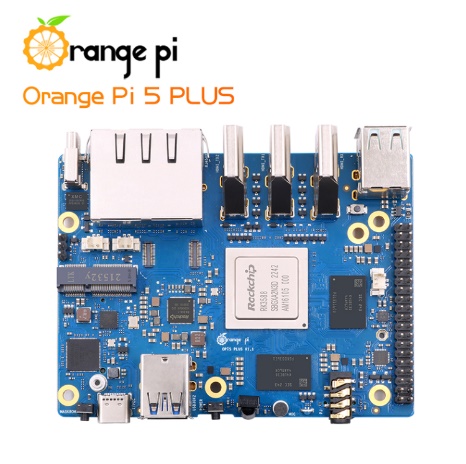
Chính xác cao: Công nghệ xử lý ảnh của OpenCV đảm bảo nhận diện chính xác cả trong điều kiện ánh sáng yếu hoặc mã vạch mờ.

Chi phí thấp: Sử dụng camera thông thường và thư viện tích hợp sẵn nguồn mở giúp tiết kiệm chi phí triển khai.

**3. Phần cứng**

**3.1. Các thành phần chính**

**3.1.1. Máy tính nhúng**

Yêu cầu: Máy tính nhỏ gọn, hiệu năng đủ để xử lý các tác vụ như nhận diện mã vạch, gợi ý sách, giao tiếp bằng giọng nói, và tích hợp các thư viện như OpenCV hoặc LangChain.

Giải pháp: Orange Pi 5 plus:

CPU ARM hiệu suất cao, hỗ trợ các ứng dụng AI/ML.

Tích hợp các cổng giao tiếp như USB, HDMI, và GPIO, dễ dàng kết nối với các thiết bị ngoại vi.

Hỗ trợ các hệ điều hành phổ biến và thư viện mã nguồn mở.

Tiết kiệm năng lượng: Công suất tiêu thụ tối thiểu, hoạt động ổn định trong thời gian dài.

**3.1.2. Camera phân giải cao**

Yêu cầu: Hỗ trợ nhận diện mã vạch chính xác ngay cả trong điều kiện ánh sáng yếu. Độ phân giải tối thiểu 2MP để đảm bảo chất lượng hình ảnh tốt.

Giải pháp:

Camera webcam 1080p

Tích hợp dễ dàng với máy tính nhúng qua cổng USB.

Hỗ trợ điều chỉnh tiêu cự và xử lý hình ảnh tốc độ cao.

Tích hợp Microphone nhạy và chính xác, thu âm rõ ràng từ khoảng cách xa, có khả năng lọc tạp âm để hoạt động tốt trong môi trường tạp âm nhẹ.

**3.1.3. Loa phát âm thanh**

Yêu cầu:

Âm thanh rõ ràng để giao tiếp với người dùng, hỗ trợ giọng nói tự nhiên của Google TTS.

Công suất vừa đủ để sử dụng trong không gian thư viện.

Hiệu suất âm thanh ổn định, chi phí hợp lý.

Tích hợp hệ thống chống nhiễu để đảm bảo âm thanh không bị méo.

**3.2. Thiết kế, in 3D và gia công vỏ robot**

Vỏ robot thư viện đóng vai trò bảo vệ các linh kiện phần cứng, đồng thời tạo nên diện mạo gần gũi và thân thiện với người dùng. Để đạt được hiệu quả cả về mặt kỹ thuật lẫn thẩm mỹ, thiết kế và gia công vỏ robot cần đáp ứng các tiêu chí sau:

* Gần gũi và thân thiện:

Hình dáng đơn giản, mềm mại với các góc bo tròn, màu sắc tươi sáng để phù hợp với không gian thư viện và dễ dàng thu hút học sinh, sinh viên.

Có màn hình cảm ứng hiển thị giao diện, tích hợp camera ở vị trí mắt để tạo cảm giác như một “người bạn” công nghệ.

* Chi phí thấp:

Sử dụng vật liệu giá rẻ nhưng bền, dễ gia công như nhựa PLA, ABS hoặc mica.

Gia công bằng công nghệ in 3D hoặc cắt CNC để tiết kiệm thời gian và chi phí.

* Nhỏ gọn và tối ưu không gian:

Kích thước phù hợp với thư viện, không chiếm nhiều diện tích nhưng vẫn đủ để chứa các linh kiện phần cứng như Orange Pi, camera, loa, và nguồn điện.

* Dễ tháo lắp và bảo trì:

Vỏ có các khớp nối hoặc vít để dễ dàng mở ra khi cần sửa chữa hoặc nâng cấp linh kiện.

**3.2.1. Quy trình thiết kế 3D**

Phần mềm thiết kế: Sử dụng các phần mềm như Fusion 360 hoặc SolidWorks để tạo mô hình 3D chính xác.

Thiết kế cấu trúc vỏ:

Thân chính: Bảo vệ phần cứng, tích hợp được màn hình và loa.

Đế robot: Đảm bảo độ ổn định khi đặt cố định trên mặt phẳng.

**3.2.1.1 Nguyên mẫu 3D**

Để tối ưu hóa trải nghiệm sử dụng và đảm bảo hiệu quả hoạt động của robot thư viện, chúng tôi đã xây dựng hai nguyên mẫu với cách bố trí camera và màn hình khác nhau.

**a. Nguyên mẫu 1**

*Bố trí:*

Camera: Được đặt trên phần đầu robot, giúp quan sát bao quát không gian và dễ dàng quét sách từ trên cao.

Màn hình: Đặt trên thân robot, cố định ở vị trí ngang tầm mắt người sử dụng khi ngồi.

*Ưu điểm:*

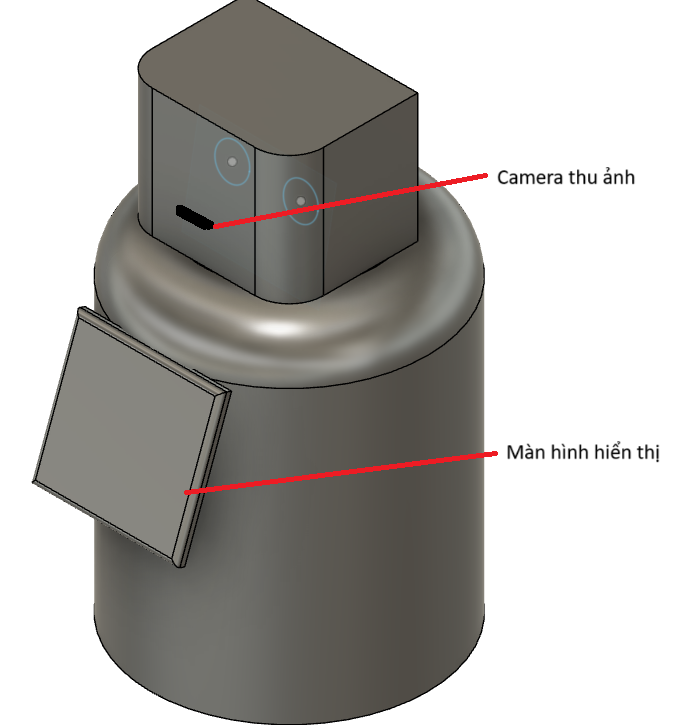
Camera trên đầu có thể nhận diện và quét mã vạch ở nhiều góc độ mà không cần cúi xuống.

Thiết kế trực quan, thân thiện với người dùng vì màn hình luôn hướng về phía người sử dụng.

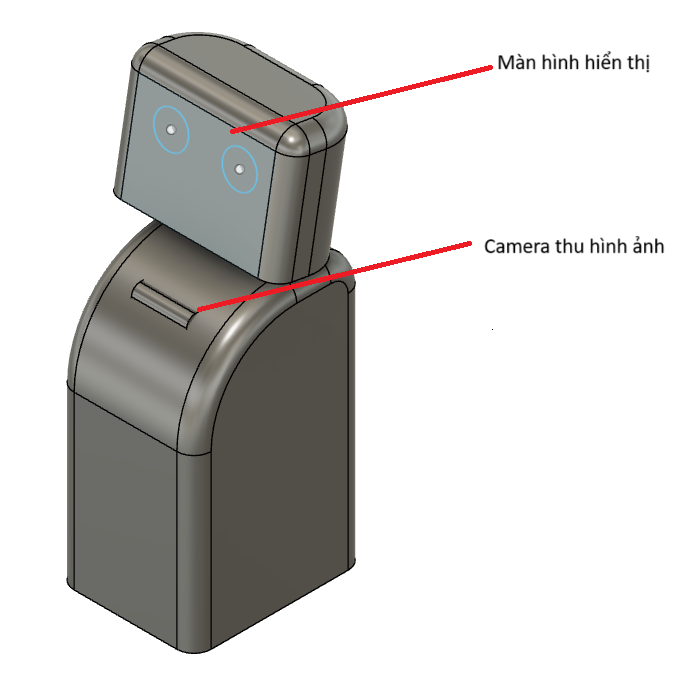
*Nhược điểm:*

Khi người dùng đưa sách lên gần camera để quét, tay và sách có thể che mất màn hình, gây khó khăn trong việc kiểm tra trạng thái nhận diện.

Màn hình cố định trên thân, góc nhìn hạn chế, khó điều chỉnh theo chiều cao của người sử dụng.

****

**b. Nguyên mẫu 2**

****

*Bố trí:*

Camera: Được đặt trên phần ngực robot, ở vị trí ngang tầm với sách, giúp dễ dàng quét mã vạch mà không che khuất màn hình.

Màn hình: Đặt trên đầu robot, vừa đóng vai trò hiển thị thông tin, vừa thể hiện biểu cảm, tạo cảm giác thân thiện.

*Ưu điểm:*

Màn hình trên đầu có thể nghiêng và xoay, giúp người dùng dễ dàng quan sát ở nhiều góc độ.

Camera ở ngực tối ưu cho việc quét mã vạch mà không lo bị che khuất.

Màn hình trên đầu tạo sự hấp dẫn hơn với khả năng hiển thị biểu cảm như mắt, miệng, tăng tính tương tác.

*Nhược điểm:*

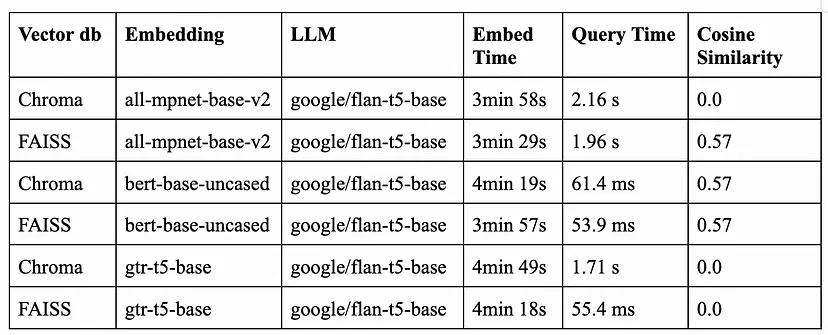
Camera ở ngực có thể bị giới hạn trong việc quét các vật thể ở khoảng cách xa, do góc nhìn hẹp hơn so với đặt trên đầu.

Thiết kế phức tạp hơn, yêu cầu cấu trúc cơ học để hỗ trợ màn hình xoay và nghiêng linh hoạt.

Nhóm đã lựa chọn nguyên mẫu 2 với ưu thế về tính tương tác, tính thân thiện, và khả năng quan sát tối ưu, nguyên mẫu 2 phù hợp hơn để triển khai trong môi trường thư viện. Nó không chỉ cải thiện trải nghiệm của người dùng mà còn tăng sức hút và sự gần gũi của robot.

**4. Thử nghiệm và đánh giá hệ thống**

**4.1. Thử nghiệm tính năng gợi ý sách**

****

Bảng 1: Bảng kết quả thử nghiệm khả năng gợi ý sách

Embedding: Mô hình bert-base-uncased cho kết quả tốt hơn trong việc tạo ra các vector đại diện cho ý nghĩa của văn bản và cải thiện độ chính xác của việc gợi ý sách.

Vector database: Cả Chroma và FAISS đều có thể được sử dụng, nhưng FAISS thường có tốc độ truy vấn nhanh hơn.

**4.2. Thử nghiệm tính năng quét mã vạch sách**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Môi trường | Điều kiện | Thử nghiệm | Số lần thử nghiệm | Kết quả trung bình |
| Ánh sáng tốt | Thư viện có ánh sáng tự nhiên hoặc ánh sáng đèn LED ổn định | Quét mã vạch sách từ các khoảng cách (10cm, 20cm, 30cm). | 10 | - Tốc độ quét < 2 giây  - Tỷ lệ nhận diện chính xác > 95% |
| Ánh sáng yếu | Môi trường ánh sáng đèn mờ, thiếu sáng | Quét mã vạch sách từ các góc nghiêng khác nhau (0°, 30°, 45°). | 10 | - Tốc độ quét < 5 giây  - Tỷ lệ nhận diện chính xác > 85% |
| Nhiễu sáng | Thư viện có ánh sáng thay đổi liên tục (ánh sáng nhấp nháy) | Quét mã vạch sách có màu sắc khác nhau (đen trắng, màu). | 10 | - Tốc độ quét < 3 giây  - Tỷ lệ nhận diện chính xác > 90% |

Bảng 2: Bảng kết quả thử nghiệm khả năng nhận diện mã vạch sách

Kết luận: Robot đạt hiệu suất nhận diện mã vạch khá tốt trong hầu hết các môi trường thử nghiệm.

**4.3. Thử nghiệm tính năng giao tiếp giọng nói**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Môi trường | Điều kiện | Thử nghiệm | Số lần thử nghiệm | Kết quả trung bình |
| Yên tĩnh | Thư viện không có tiếng ồn | Người dùng yêu cầu thông tin sách qua giọng nói với âm lượng bình thường. | 10 | - Tỷ lệ nhận diện đúng lời nói > 98% |
| Nhiễu nhẹ | Môi trường có tiếng thì thầm hoặc máy lạnh hoạt động | Người dùng yêu cầu thông tin sách bằng giọng nói nhỏ. | 10 | - Tỷ lệ nhận diện đúng lời nói > 90% |
| Nhiễu trung bình | Thư viện có tiếng nói chuyện từ xa hoặc tiếng bước chân | Người dùng yêu cầu thông tin sách với các tốc độ nói khác nhau (chậm, trung bình, nhanh). | 10 | - Tỷ lệ nhận diện đúng lời nói > 85% |
| Nhiễu cao | Thư viện có tiếng ồn lớn từ máy móc hoặc hoạt động cải tạo | Người dùng yêu cầu thông tin sách bằng giọng nói to và rõ. | 10 | - Tỷ lệ nhận diện đúng lời nói > 70% |
| Nhiều người nói cùng lúc | Môi trường có nhiều người nói cùng lúc gần robot | Robot nhận diện và trả lời giọng nói chính xác với người giao tiếp. | 10 | - Độ chính xác trong phân biệt giọng nói > 80% |

Kết luận: Robot có thể duy trì độ chính xác cao trong môi trường yên tĩnh và nhiễu nhẹ. Với nhiễu trung bình hoặc cao, cần tích hợp thuật toán giảm nhiễu hoặc khuyến nghị chuyển sang giao tiếp bằng màn hình cảm ứng để tránh sai sót đáng kể.

**IV. Kết luận**

**1. Đánh giá kết quả:**

Robot được thiết kế nhỏ gọn, tiết kiệm năng lượng, đảm bảo tính gần gũi với người dùng. Cấu hình phần cứng phù hợp với nhu cầu xử lý các tác vụ phức tạp, đồng thời duy trì chi phí thấp để phù hợp với môi trường trường học. Robot vận hành ổn định trong các môi trường thử nghiệm khác nhau, từ thư viện có điều kiện ánh sáng lý tưởng đến môi trường nhiều tiếng ồn. Robot giúp người dùng dễ dàng tìm kiếm, mượn trả sách và nhận gợi ý sách cá nhân hóa, hỗ trợ thư viện hiện đại hóa dịch vụ. Cải thiện quản lý sách: Hệ thống nhận diện và lưu trữ thông tin hiệu quả, giảm sai sót và thời gian xử lý thủ công.

**2. Hướng phát triển**

Camera hỗ trợ ánh sáng yếu: Bổ sung đèn LED tự động bật khi điều kiện ánh sáng kém.

Thuật toán giảm nhiễu âm thanh: Áp dụng kỹ thuật lọc tiếng ồn hoặc beamforming để cải thiện khả năng nhận diện giọng nói.

Hỗ trợ đa ngôn ngữ: Đảm bảo robot có thể nhận diện nhiều ngôn ngữ và phương ngữ để phục vụ đa dạng người dùng.

Tăng cường thử nghiệm thực tế: Mời người dùng thử nghiệm trực tiếp trong các thư viện thật để tinh chỉnh hiệu suất theo phản hồi thực tế.

**TÀI LIỆU THAM KHẢO**

[1] [(PDF) Thiết Kế và Điều Khiển Mô Hình Robot Thư Viện](https://www.researchgate.net/publication/385340515_Thiet_Ke_va_Dieu_Khien_Mo_Hinh_Robot_Thu_Vien)

[2] Hướng dẫn sử dụng Orange Pi 5 Plus: <http://www.orangepi.org/orangepiwiki/index.php/Orange_Pi_5>

[3] Nguyên lý hoạt động mô hình LLM:

<https://github.com/mlabonne/llm-course>

[4] Bộ dữ liệu sách tiki

[Tiki Books Dataset](https://www.kaggle.com/datasets/biminhc/tiki-books-dataset)