**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**KHOA HỆ THỐNG THÔNG TIN**

**NGUYỄN ĐỨC TÙNG – 16521396**

**HÀ THANH ĐOÀN - 16521396**

**XÂY DỰNG HỆ THỐNG GỢI Ý CHO**

**WEBSITE REVIEW NHÀ HÀNG/QUÁN ĂN**

**RECOMMENDER SYSTEM FOR**

**RESTAURANTS/EATERIES REVIEWING WEBSITE**

**KỸ SƯ NGÀNH HỆ THỐNG THÔNG TIN**

**GIẢNG VIÊN HƯỚNG DẪN**

**THS. THÁI BẢO TRÂN**

**TP. HỒ CHÍ MINH, 12/2020**

DANH SÁCH HỘI ĐỒNG BẢO VỆ KHÓA LUẬN

Hội đồng chấm khóa luận tốt nghiệp, thành lập theo Quyết định số …………………… ngày ………………….. của Hiệu trưởng Trường Đại học Công nghệ Thông tin.

* 1. …………………………………………. – Chủ tịch.
  2. …………………………………………. – Thư ký.
  3. …………………………………………. – Ủy viên.
  4. …………………………………………. – Ủy viên.

|  |  |
| --- | --- |
| **ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**  **TRƯỜNG ĐẠI HỌC**  **CÔNG NGHỆ THÔNG TIN** | **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**  **Độc Lập - Tự Do - Hạnh Phúc** |
|  | *TP. HCM, ngày…..tháng…..năm……..* |

**NHẬN XÉT KHÓA LUẬN TỐT NGHIỆP**

**(CỦA CÁN BỘ HƯỚNG DẪN)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Tên khóa luận:** | | | |
| **ĐỌC VÀ TƯ VẤN SÁCH ĐIỆN TỬ TRÊN ỨNG DỤNG DI ĐỘNG** | | | |
| **Nhóm SV thực hiện:** | | | **Cán bộ hướng dẫn:** |
| Nguyễn Đức Tùng  Hà Thanh Đoàn | 16521396  16520219 | | ThS. Thái Bảo Trân |
| **Đánh Giá Khóa Luận:**   1. **Về cuốn báo cáo:**   Số trang Số chương  Số bảng số liệu Số hình vẽ  Số tài liệu tham khảo Sản phẩm  **Một số nhận xét về hình thức cuốn báo cáo:**   1. **Về nội dung nghiên cứu:** 2. **Về chương trình ứng dụng:** 3. **Về thái độ làm việc của sinh viên:**   **Đánh giá chung:**  **Điểm từng sinh viên:**  **Nguyễn Đức Tùng:**  **Hà Thanh Đoàn:** | | | |
|  | | **Người nhận xét** | | |

|  |  |
| --- | --- |
| **ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**  **TRƯỜNG ĐẠI HỌC**  **CÔNG NGHỆ THÔNG TIN** | **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**  **Độc Lập - Tự Do - Hạnh Phúc** |
|  | *TP. HCM, ngày…..tháng…..năm……..* |

**NHẬN XÉT KHÓA LUẬN TỐT NGHIỆP**

**(CỦA CÁN BỘ PHẢN BIỆN)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Tên khóa luận:** | | | |
| **XÂY DỰNG HỆ THỐNG GỢI Ý CHO WEBSITE REVIEW NHÀ HÀNG/QUÁN ĂN** | | | |
| **Nhóm SV thực hiện:** | | | **Cán bộ phản biện:** |
| Nguyễn Đức Tùng  Hà Thanh Đoàn | 16521396  16520219 | |  |
| **Đánh giá Khóa luận**   1. Về cuốn báo cáo:   Số trang Số chương  Số bảng số liệu Số hình vẽ  Số tài liệu tham khảo Sản phẩm  **Một số nhận xét về hình thức cuốn báo cáo:**   1. **Về nội dung nghiên cứu:** 2. **Về chương trình ứng dụng:** 3. **Về thái độ làm việc của sinh viên:**   **Đánh giá chung:**  **Điểm từng sinh viên:**  **Nguyễn Đức Tùng:**  **Hà Thanh Đoàn:** | | | |
|  | | **Người nhận xét** | | |
|  | |  | | |

# LỜI CẢM ƠN

Chúng em xin gửi lời cảm ơn sâu sắc đến cô Thái Bảo Trân. Cô là người đã hướng dẫn, chỉ lối và gợi ý để chúng em có thể để hoàn thành khóa luận này. Cô luôn nhiệt tình liên lạc, trao đổi và tận tình góp ý giúp chúng em hoàn thành khóa luận với kết quả tốt nhất.

Chúng em xin chân thành biết ơn sự tận tình dạy dỗ và sự giúp đỡ của tất cả quý thầy cô tại trường Đại học Công Nghệ Thông Tin, đặc biệt là các thầy cô trong khoa Hệ Thống Thông Tin, các thầy cô đã tận tình truyền đạt những kiến thức bổ ích, quý giá cho chúng em trong suốt hơn 4 năm học qua, những kiến thức học được từ giảng đường sẽ là hành trang quý báu để chúng em có thể tiếp tục trên con đường học tập, nghiên cứu, làm việc cũng nhưng trong quá trình hoàn thiện bản thân của mình.

Chúng em xin chân thành cảm ơn quý thầy cô!

Xin trân trọng cảm ơn!

|  |  |
| --- | --- |
| **ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**  **TRƯỜNG ĐẠI HỌC**  **CÔNG NGHỆ THÔNG TIN** | **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**  **Độc Lập - Tự Do - Hạnh Phúc** |

# ĐỀ CƯƠNG CHI TIẾT

|  |  |
| --- | --- |
| **TÊN ĐỀ TÀI:**  **Xây dựng hệ thống gợi ý cho website review nhà hàng/quán ăn**  **(Recommender system for restaurants/eateries reviewing website)** | |
| **Cán bộ hướng dẫn:** ThS. Thái Bảo Trân | |
| **Thời gian thực hiện:** Từ 9/2020 đến 12/2020 | |
| **Sinh viên thực hiện:**  Nguyễn Đức Tùng - 16521396  Hà Thanh Đoàn - 16520219 | |
| **NỘI DUNG ĐỀ TÀI**   1. **Lý do chọn đề tài**  * Nhu cầu ăn uống là một trong những nhu cầu cần thiết trong cuộc sống. Thế nhưng rất khó để tìm quán ăn phù hợp với sở thích về nhiều yếu tố như quang cảnh, chất lượng, giá cả,… Đặc biệt đối với khách du lịch, việc tìm kiếm một nhà hàng/quán ăn phù hợp luôn là một trong những điều được quan tâm hàng đầu. * Ăn gì, ở đâu? Là bài toán không còn mới. Tuy nhiên, với một số website review nhà hàng hiện nay, người dùng phải tìm kiếm từng nhà hàng, xem mô tả, đọc các bình luận,.. Điều này thật sự tốn thời gian. Hơn nữa, việc chỉ dựa vào các đánh giá (chấm điểm, bình luận) của những khách hàng trước đó sẽ không hoàn toàn phù hợp với nhu cầu vì mỗi người có một sở thích, khẩu vị riêng. Do đó, việc webiste tự động đề xuất những nhà hàng/quán ăn phụ hợp với người dùng hoặc gợi ý những nhà hàng tương tự sẽ giúp người dùng dễ dàng tiếp cận những nơi phù hợp với cá nhân. * Qua khảo sát các website trong nước, chúng tôi nhận thấy hiện nay chưa có website nào xây dựng được chức năng đề xuất. Đa số các website đơn giản chỉ hiển thị các nhà hàng gần/mới nhất mà không quan tâm đến nhu cầu của người dùng. Vì vậy, chúng tôi chọn đề tài này với mong muốn thiết kế một trang web, ngoài những chức năng của một website review nhà hàng, hệ thống tập trung giải quyết thêm vấn đề khuyến nghị cho người dùng có thể tìm kiếm quán ăn phù hợp với sở thích của mình nhất.  1. **Mục tiêu đề tài**  * Nhóm người dùng hướng đến: Dành cho tất cả mọi người, đặc biệt giới trẻ. * Áp dụng các kỹ thuật trong khai thác dữ liệu để xây dựng hệ thống gợi ý cho website, nhằm mang lại những đánh giá đúng nhất về các nhà hàng dựa trên trải nghiệm người dùng đã chia sẻ. * Ngoài ra, để tạo ra sự thuận tiện cho người dùng, chúng tôi sẽ tích hợp thêm vào hệ thống chức năng tìm kiếm bằng giọng nói.  1. **Phạm vi đề tài**   Đề tài tập trung vào các nội dung chính sau:   * Các chức năng cơ bản của website (Hiển thị, quản lý, thống kê,…) * Hệ thống gợi ý nhà hàng phù hợp nhất với người dùng * Tìm kiếm bằng giọng nói. * Tư vấn các nhà hàng có khoảng cách gần người dùng nhất. * Lưu lịch sử thao tác người dùng như: những nhà hàng mà người dùng vào xem, thời gian xem,… * Hiện tại chúng tôi chỉ xây dựng hệ thống gợi ý dựa trên các dữ liệu đánh giá thu thập được từ khách hàng đã trãi nghiệm dịch vụ. Do đó, kết quả hệ thống gợi ý sẽ không khai thác tối đa lợi ích đối với địa điểm không có hoặc có ít thông tin khách hàng đánh giá sau khi họ trải nghiệm.  1. **Đối tượng**  * Tất cả mọi người.  1. **Phương pháp thực hiện**  * Khảo sát và phân tích nhu cầu của người dùng. * Thiết kế lược đồ Cơ Sở Dữ Liệu (CSDL) và xây dựng quy trình nghiệp vụ. * Tìm hiểu các công nghệ áp dụng để xây dựng các chức năng phù hợp. * Ứng dụng các thuật toán phù hợp để xây dựng hệ thống gợi ý. * Xây dựng chức năng tìm kiếm giọng nói bằng một số kỹ thuật xử lý ngôn ngữ tự nhiên.  1. **Kết quả mong đợi của đề tài**  * Website trực quan, dễ sử dụng, giao diện thân thiện với mọi thiết bị. * Xây dựng được hệ thống quản lý website (admin) * Hệ thống gợi ý và tìm kiếm giọng nói với độ chính xác cao. * Đề tài có tính ứng dụng thực tiễn cao. Đáp ứng nhu cầu của người dùng. | |
| **KẾ HOẠCH THỰC HIỆN** | |
| **Công việc cần thực hiện** | **Thời gian** |
| * Khảo sát hiện trạng * Viết mô tả cho đề tài * Tìm hiểu các công nghệ liên quan áp dụng cho đề tài. | 1 tuần |
| * Thu thập dữ liệu. * Tìm hiểu và thiết kế CSDL cho hệ thống | 1 tuần |
| * Tìm hiểu các thuật toán, kỹ thuật và hiện thực với dữ liệu có sẵn. | 3 tuần |
| * Thiết kế giao diện cho webiste | 2 tuần |
| * Lập trình backend. | 3 tuần |
| * Kiểm tra và sửa lỗi phát sinh. Cải thiện trang web | 2 tuần |
| * Triển khai hệ thống | 1 tuần |
| * Hoàn chỉnh báo cáo đề tài | 2 tuần |
| **Tổng cộng** | **15 tuần** |
| **Xác nhận của CBHD**  (Ký tên và ghi rõ họ tên)  **Thái Bảo Trân** | **TP. HCM, 26 / 07 / 2020**  **Sinh viên**  (Ký tên và ghi rõ họ tên) |

**MỤC LỤC**

[LỜI CẢM ƠN 7](#_Toc58013729)

[ĐỀ CƯƠNG CHI TIẾT 8](#_Toc58013730)

[DANH MỤC CÁC BẢNG 13](#_Toc58013731)

[DANH MỤC CÁC HÌNH 14](#_Toc58013732)

[DANH MỤC THUẬT NGỮ VÀ TỪ VIẾT TẮT 15](#_Toc58013733)

[TÓM TẮT 17](#_Toc58013734)

[CHƯƠNG 1: TỔNG QUAN 1](#_Toc58013735)

[1.1. Đặt vấn đề 1](#_Toc58013736)

[1.2. Mục tiêu đề tài 1](#_Toc58013737)

[1.3. Khảo sát hiện trạng 2](#_Toc58013738)

[1.3.1. Khảo sát website review nhà hàng/quán ăn 2](#_Toc58013739)

[1.3.2. Các nghiên cứu liên quan 3](#_Toc58013740)

[1.4. Đối tượng, phạm vi nghiên cứu 4](#_Toc58013741)

[1.4.1. Đối tượng 4](#_Toc58013742)

[1.4.2. Phạm vi nghiên cứu 4](#_Toc58013743)

[1.5. Nội dung thực hiện 4](#_Toc58013744)

[1.6. Bố cục báo cáo 5](#_Toc58013745)

[CHƯƠNG 2: CƠ SỞ LÝ THUYẾT 7](#_Toc58013746)

[2.1. Tổng quan về cơ sở dữ liệu MySQL 7](#_Toc58013747)

[2.1.1. Giới thiệu về MySQL 7](#_Toc58013748)

[2.1.2. Cấu trúc lưu trữ dữ liệu 7](#_Toc58013749)

[2.2. Phương thức hoạt đông của MySQL 7](#_Toc58013750)

[2.3. Các chức năng MySQL hỗ trợ 7](#_Toc58013751)

[2.4. Thuật toán Lọc cộng tác(Collaborative Filtering) 7](#_Toc58013752)

[2.5. Thuật toán Điều khiển bằng giọng nói 7](#_Toc58013753)

[2.6. Chức năng tìm kiếm các nhà hàng/quán ăn trên bản đồ 7](#_Toc58013754)

[2.7. Các công cụ hỗ trợ hệ thống 7](#_Toc58013755)

[2.7.1. Python 7](#_Toc58013756)

[2.7.2. Php 7](#_Toc58013757)

[CHƯƠNG 3: PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG 8](#_Toc58013758)

[3.1. Phân tích hiện trạng nhu cầu 8](#_Toc58013759)

[3.1.1. Yêu cầu chức năng 8](#_Toc58013760)

[3.1.2. Yêu cầu phi chức năng: 8](#_Toc58013761)

[CHƯƠNG 4: CÀI ĐẶT VÀ TRIỂN KHAI HỆ THỐNG 9](#_Toc58013762)

[CHƯƠNG 5: KẾT LUẬN 10](#_Toc58013763)

[TÀI LIỆU THAM KHẢO 11](#_Toc58013764)

DANH MỤC CÁC BẢNG

DANH MỤC CÁC HÌNH

# DANH MỤC THUẬT NGỮ VÀ TỪ VIẾT TẮT

|  |  |  |
| --- | --- | --- |
| **STT** | **Từ viết tắt** | **Nội dung** |
| 1 |  |  |
| 2 |  |  |
| 3 |  |  |

TÓM TẮT

Đề tài khóa luận “Xây dựng hệ thống gợi ý cho website review nhà hàng/quán ăn” mang tính thực tế, hướng đến những đối tượng có nhu cầu tìm các nhà hàng/quán ăn hợp với sở thích hoặc giá tiền nhưng cần các nhận xét hoặc gợi ý chân thực đến từ những người đã từng trải nghiệm các nhà hàng này.

Sau khi tìm hiểu, khảo sát các ứng dụng review nhà hàng/quán ăn, nhóm tác giả đưa ra hướng giải quyết sau:

* Tìm hiểu và xây dựng phương pháp khuyến nghị dựa vào lịch sử nhận xét hoặc vị trí của người dùng, đồng thời tích hợp điều khiển bằng giọng nói.
* Xây dựng ứng dụng với chức năng tìm kiếm và đọc review với giao diện thân thiện, dễ sử dụng. Ngoài ra, ứng dụng còn cung cấp các chức năng khác như: đăng nhập, đăng xuất, tìm kiếm, đánh giá nhà hàng, chia sẻ nhà hàng, tạo nhà hàng.
* Công nghệ sử dụng xây dựng hệ thống: Python, PHP(Laravel Framework), Heroku Cloud Platform.

# CHƯƠNG 1: TỔNG QUAN

* 1. **Đặt vấn đề**

Ngay từ khi những dấu chân đầu tiên của con người in dấu lên Trái Đất, ngay từ thời kỳ bình minh của nhân loại, việc ăn uống đã được coi như là một nhu cầu cần thiết để duy trì sự sống của sinh vật nói chung và con người nói riêng. Tuy nhiên, ở thời kỳ cổ đại đó, thức ăn vẫn còn khan hiếm và đơn giản, buộc con người không có quyền lựa chọn thức ăn. Sau này, trải qua hàng triệu năm tiến hóa, thế giới ngày một văn minh hơn, do đó, những tri thức cơ bản đầu tiên về lĩnh vực ăn uống được hình thành, tạo nên khái niệm đầu tiên về văn hóa ăn uống: Ẩm thực.

[Theo thống kê của Dcorp R-Keeper Việt Nam](https://vietnambiz.vn/nganh-fb-viet-nam-2025-co-the-dat-408-ti-usd-doanh-thu-201912071250533.htm) vào năm 2019 cho biết cả nước hiện có đến 540.000 cửa hàng ăn uống, trong đó có khoảng 430.000 cửa hàng nhỏ, 7.000 nhà hàng chuyên dịch vụ thức ăn nhanh, 22.000 cửa hàng cà phê, các quầy bar và trên 80.000 nhà hàng được đầu tư, phát triển một cách bài bản. Qua từng năm, con số này tăng lên một cách nhanh chóng và đáng kể.

Xã hội càng văn minh, càng hiện đại, con người lại càng có góc nhìn và nhu cầu khắt khe hơn về ẩm thực, không còn gói gọn như câu nói ngày xưa của ông bà ta “Ăn no mặc ấm” mà bây giờ đã chuyển thành “Ăn ngon mặc đẹp”.

[Theo thống kê của Vnetwork](https://vnetwork.vn/news/thong-ke-internet-viet-nam-2020), tính đến tháng 1 năm 2020, có khoảng 68,17 triệu người Việt Nam sử dụng internet, chiếm khoản 70% dân số cả nước. Điều này chứng tỏ rằng người Việt Nam sử dụng internet trong cuộc sống rất nhiều. Và tất nhiên việc tìm kiếm nhà hàng/quán ăn bằng internet cũng không ngoại lệ.

Với sự phát triển của công nghệ thông tin, việc tìm kiếm nhà hàng/quán ăn sẽ dễ dàng hơn cho mọi người vì chỉ cần lên website và tra trên Google. Nhưng rất khó tìm được những trang uy tín và có nhận xét chân thực.

Nhằm hỗ trợ cho việc tìm kiếm nhà hàng/quán ăn, đặc biệt dành cho người Việt, nhóm tác giả cố gắng tạo ra một phương pháp giúp người dùng có thể tìm kiếm nhà hàng/quán ăn phù hợp với mình, đồng thời nhận xét và đánh giá các nhà hàng mà chính người dùng đã trải nghiệm. Từ đó tạo nên một cộng đồng chia sẻ, đánh giá nhà hàng/quán ăn đáng tin cậy.

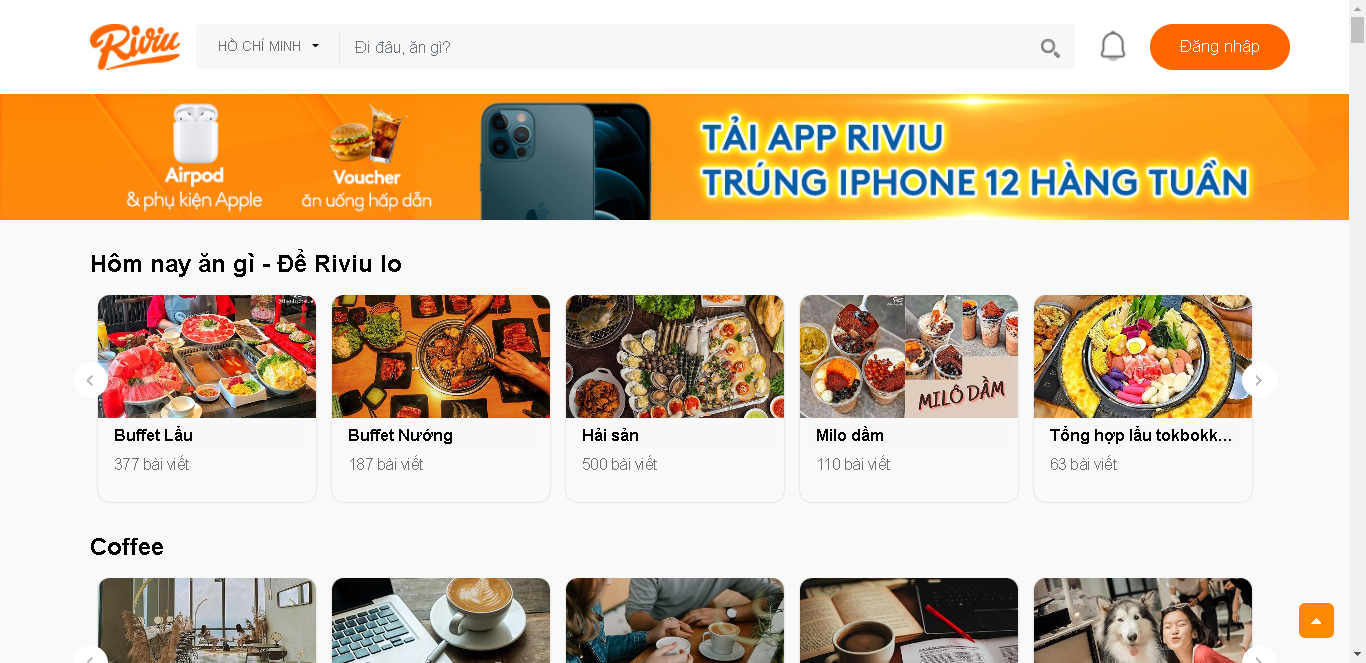
* 1. **Mục tiêu đề tài**

Mặc dù đã có rất nhiều trang quảng cáo hoặc nhận xét về chất lượng cũng như giá cả của nhà hàng/quán ăn, nhưng hầu hết lại khá rời rạc và hoàn toàn không thuần túy là trang review nhà hàng/quán ăn. Vì vậy trang web này sẽ giúp người dùng tìm được nhà hàng/quán ăn hợp lý cũng như đưa ra các nhận xét dựa theo trải nghiệm của bản thân.

Để tiện lợi cho việc tìm kiếm, nhóm hướng tới một website review nhà hàng/quán ăn để giúp người dùng có thể dễ dàng truy cập trên máy tính, laptop hoặc cả trên điện thoại. Website giúp người dùng dễ dàng sử dụng để tìm kiếm nhà hàng/ quán ăn mà mình cần đến, những quán ăn xung quanh mình hoặc đưa ra những thang điểm, nhận xét về nhà hàng mình vừa ăn. Bên cạnh đó, ứng dụng tạo nên một cộng đồng cùng nhau bàn luận cũng như gợi ý về các nhà hàng/quán ăn để tiện lợi hơn cho việc chọn lựa.

* 1. **Khảo sát hiện trạng**
     1. **Khảo sát website review nhà hàng/quán ăn**
* [riviu.vn](https://riviu.vn/)

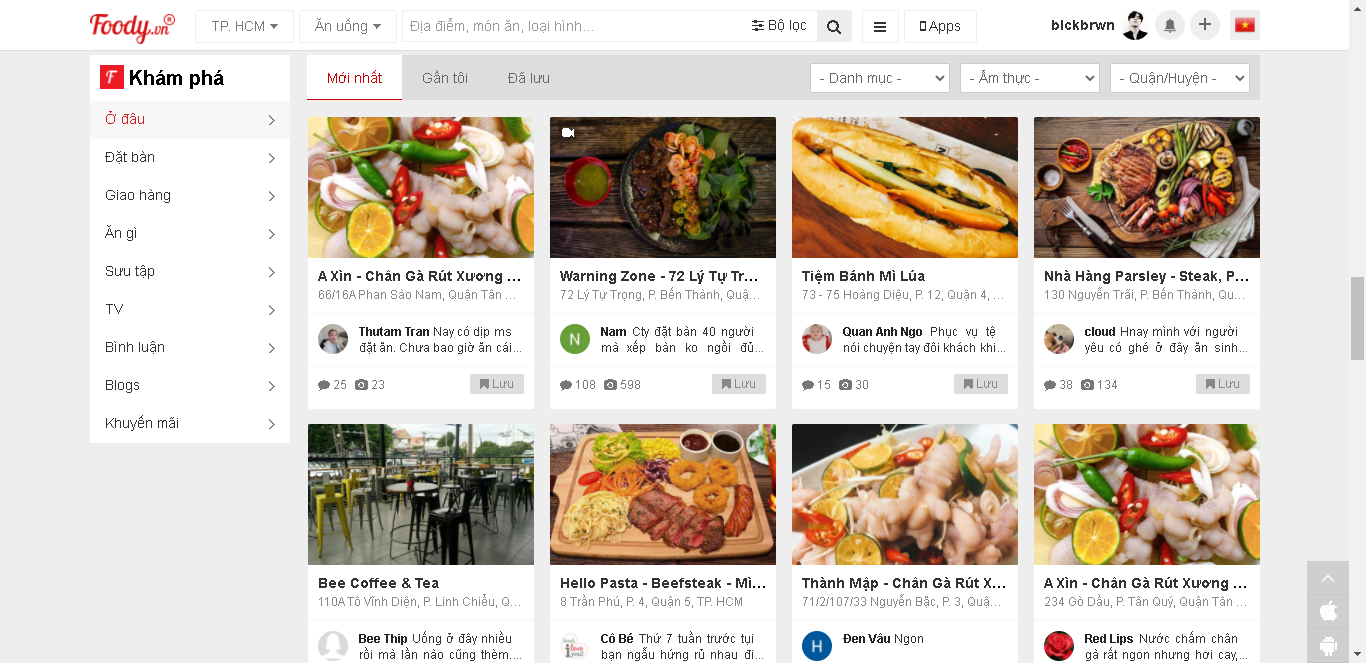
Đúng như tên gọi, riviu.vn là website ẩm thực, du lịch hàng đầu Việt Nam. Là diễn đàn để cộng đồng thảo luận, chia sẻ những bài review về những hàng quán ngon tuyệt của dân “sành ăn” và những kinh nghiệm du lịch tuyệt vời đã được cộng đồng trải nghiệm ở khắp mọi miền đất nước.



* [foody.vn](https://www.foody.vn/#!)

Được xây dựng từ giữa năm 2012 tại TP. HCM, Việt Nam, Foody là cộng đồng tin cậy cho mọi người có thể tìm kiếm, đánh giá, bình luận các địa điểm ăn uống: nhà hàng, quán ăn, cafe, bar, karaoke, tiệm bánh, khu du lịch… tại Việt Nam, từ website hoặc ứng dụng di động. Tất cả thành viên từ Bắc đến Nam, Foody kết nối những thực khách đến với các địa điểm ăn uống lớn nhỏ cả đất nước.

Đến thời điểm hiện tại, Foody với hàng chục ngàn địa điểm và hàng ngàn bình luận, hình ảnh tại Việt Nam ở hầu hết các tỉnh thành. Foody là cách dễ nhất để bạn có thể tìm kiếm và lựa chọn địa điểm tốt nhất cho mình và bạn bè.



* Nhận xét: Nhìn chung, chúng ta thấy được các website review nhà hàng/quán ăn hiện nay đều có chức năng chính là hiển thị các nhận xét cũng như số điểm từ người dùng chấm cho các quán ăn. Tuy nhiên, việc tìm kiếm vẫn hoàn toàn phụ thuộc vào người dùng và chưa có bất kỳ một sự gợi ý tự động nào đến từ phía website. Đối với những người dùng không quyết đoán trong việc nên đi đến quán ăn nào thì việc gợi ý khá là quan trọng. Đồng thời, các website trên hoàn toàn chưa có chức năng điều khiển bằng giọng nói. Việc sử dụng hệ thống gợi ý cũng như điều khiển qua giọng nói sẽ giúp người dùng dễ dàng hơn trong việc tìm kiếm cũng như đưa ra quyết định khi tìm kiếm nhà hàng/quán ăn.
  + 1. **Các nghiên cứu liên quan**
* Thuật toán khuyến nghị Lọc cộng tác(Collaborative Filtering).
* Thuật toán điều khiển bằng giọng nói(Tf-tfd).
  1. **Đối tượng, phạm vi nghiên cứu**
     1. **Đối tượng**

Mọi đối tượng có sử dụng website và có nhu cầu tìm kiếm các nhà hàng/quán ăn cũng như đưa ra nhận xét, thang điểm cho quán ăn đã trải nghiệm.

* + 1. **Phạm vi nghiên cứu**

Về công nghệ:

* Ngôn ngữ lập trình: Python, Php.
* Thư viện: Scikit learn.
* Framework: PhpStorm, PyCharm.
* Hệ quản trị cơ sở dữ liệu: MySQL
* Quản lý source code với Git.

Về nghiên cứu:

* Thuật toán khuyến nghị Lọc cộng tác(Collaborative Filtering).
* Thuật toán Điều khiển bằng giọng nói(Tf-tfd).
  1. **Nội dung thực hiện**

**Bước 1:**

* Khảo sát các website review nhà hàng/quán ăn.
* Xác định các chức năng cần có và thứ tự ưu tiên.
* Tìm hiểu các công nghệ, thư viện hỗ trợ việc lập trình ứng dụng.
* Tìm kiếm và nghiên cứu các nguồn dữ liệu.

**Bước 2:**

* Thiết kế các thành phần hệ thống.
* Thiết kế cơ sở dữ liệu.
* Thiết kế server.
* Thiết kế giao diện.
* Thực hiện lấy dữ liệu.

**Bước 3:**

* Xây dựng server.
* Tạo API tương tác với server.
* Tạo giao diện
* Tạo các chức năng đã chọn.

**Bước 4:**

* Xây dựng hệ thống khuyến nghị từ thuật toán tìm hiểu.

**Bước 5:**

* Hoàn chỉnh và cải thiện hiệu năng hệ thống.
* Nhận xét và đánh giá.
  1. **Bố cục báo cáo**
* *Chương 1:* Tổng quan về đề tài

Giới thiệu tổng quan nhất về nội dung đề tài. Nội dung chương 1 bao gồm: Đặt vấn đề và lý do chọn đề tài, khảo sát hiện trạng, mục tiêu và phạm vi nghiên cứu, công cụ xây dựng hệ thống và bố cục của báo cáo.

* *Chương 2:* Cơ sở lý thuyết

Giới thiệu lý thuyết, các thư viện, công cụ sử dụng trong đề tài. Trình bày ưu nhược điểm của từng mục và lí do lựa chọn. Trình bày chi tiết lý thuyết áp dụng trong đề tài.

* *Chương 3:* Phân tích và thiết kế hệ thống

Trình bày sơ đồ tổ chức của ứng dụng, phân tích thiết kế hệ thống, thiết kế cơ sở dữ liệu, thiết kế giao diện và luồng chạy của ứng dụng.

* *Chương 4:* Cài đặt và triển khai hệ thống

Trình bày những công cụ đã sử dụng để hiện thực sản phẩm của đề tài kèm ví dụ. Cách áp dụng các thuật toán vào bài toán đề xuất nhà hang cho người dùng.

* *Chương 5:* Kết luận

Tóm tắt đề tài, thành công và hạn chế, hướng phát triển tương lai về mặt chức năng và khả năng đề xuất cho người dùng.

# CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

## **Tổng quan về cơ sở dữ liệu MySQL**

### **Giới thiệu về MySQL**

### **Cấu trúc lưu trữ dữ liệu**

### **Phương thức hoạt đông của MySQL**

### **Các chức năng MySQL hỗ trợ**

* 1. **Công thức Cosine trong tính toán mức độ tương tự**

Công thức Cosine thích hợp cho dạng dữ liệu định lượng. Ví dụ dữ liệu điểm đánh giá của các nhà hàng như sau (thang điểm 10):

|  |  |  |
| --- | --- | --- |
|  | Nhà hàng A | Nhà hàng B |
| User 1 | 3 | 5 |
| User 2 | 4 | 1 |
| User 3 | 2 | 5 |

Công thức Cosine xem các dòng của ma trận đánh giá như một vector không gian. Từ ma trận trên ta biểu diễn trong đồ thị hai chiều như sau.

Chart

Description automatically generated

Từ những góc giữa các vector, dễ dàng nhận ra vector của user 1 và user 3 gần hơn là giữa user 1 và user 2. Từ đó có thể nhận định rằng user 1 và user 3 có sở thích (khẩu vị) giống nhau hơn.

Công thức Cosine được tính như sau:

Diagram

Description automatically generated

Tuy nhiên khi nói về việc so sánh các điểm đánh giá của người dung khác nhau, nhiều người lại sử dụng các thang điểm khác nhau khi họ đánh giá ( Một số người dễ tính cho 10/10 điểm nếu họ thích nhà hàng A, nhưng có một số người khó tính chỉ cho 7 hoặc 8/10). Do đó công thức Cosine được điều chỉnh lại như sau (1) :

Text

Description automatically generated

Trong đó:

sim( i , j ) : Độ tương tự giữa hai đối tượng i và j.

U: tập người dùng( user).

R u,i : Điểm đánh giá của user cho đối tượng i.

R­ u,j : Điểm đánh giá của user cho đối tượng j.

u: Điểm trung bình của user

sim(i,j) = [-1,1];

* 1. **Thuật toán lọc cộng tác(Collaborative Filtering):**

### **Tổng quan thuật toán**

Ý tưởng cơ bản của thuật toán lọc cộng tác là xác định mức độ quan tâm của một user tới một item dựa trên các users khác gần giống với user này. Việc gần giống nhau giữa các users có thể được xác định thông qua mức độ quan tâm của các users này tới các items khác mà hệ thống đã biết. Ví dụ, A, B đều thích nhà hàng X, tức đều đánh giá nhà hàng này 10/10 điểm. Ta đã biết A cũng thích nhà hàng Y, vậy nhiều khả năng B cũng thích nhà hàng này.

Thuật toán lọc cộng tác có thể được xử lý trong hai cách. Cách thứ nhất xác định mức độ quan tâm của một user tới một item dựa trên mức độ quan tâm của những user “có cùng khẩu vị” tới item đó ( User-based filtering). Cách thứ hai thay vì xác định user tương đồng, hệ thống sẽ xác định các item tương đồng. Từ đó hệ thống gợi ý những item gần giống những item mà user có mức độ quan tâm cao (Item-based filtering).

Diagram

Description automatically generated

### **Các bước thực thi**

#### **Cách tiếp cận User-based filtering:**

Bước 1: Xây dựng ma trận đánh giá.

Bước 2: Tính độ tương đồng giữa user hiện tại với tất cả các user còn lại bằng công thức Cosine.

Bước 3: Sắp xếp kết quả từ bước 2 giảm dần theo độ tương đồng.

Bước 4: Chọn ra những user lân cận (neighbors). Cách chọn neighborhood user được trình bày rõ ở phần sau (Cách tiếp cận Item-based filtering).

Bước 5: Dự đoán điểm đánh giá của user hiện tại với các item. Lấy ra item có điểm dự đoán cao nhất. Công thức dự đoán điểm đánh giá được trình bày rõ ở phần sau (Cách tiếp cận Item-based filtering).

Diagram

Description automatically generated

#### **Cách tiếp cận Item-based filtering:**

Bước 1: Xây dựng ma trận đánh giá.

Bước 2: Tính độ tương đồng giữa các item.

Bước 3: Sắp xếp kết quả từ bước 2 giảm dần theo độ tương đồng.

Bước 4: Chọn ra những item lân cận (neighbors). Một số cách để chọn ra neighborhood:

* Top-N: Đây là cách đơn giản để lấy ra N neighbors mà không cần quan tâm đến độ tương đồng, thậm chí lấy luôn cả những item có độ tương đồng thấp.
* Threshold: Lấy ra những item có độ tương đồng cao hơn một hằng số N cho trước.
* Ưu và nhược điểm giữa hai phương pháp:

|  |  |  |
| --- | --- | --- |
|  | **Ưu điểm** | **Nhược điểm** |
| **Top-N** | Lấy được nhiều kết quả | Vì không quan tâm đến độ tương tự nên kết quả có thể chứa những item có độ tương tự thấp, dẫn đến chất lượng gợi ý kém hiệu quả. |
| **Threshold** | Đảm bảo được kết quả luôn là những item có độ tương đồng cao ( tùy theo hằng số N cho trước). | Số lượng neighbors nhỏ. |

Chart, bubble chart

Description automatically generated

Hình bên trái là phương pháp chọn threshold, xung quanh item hiên tại vẽ vòng tròn bán kính N, những điểm nằm trong vòng tròn chính là neighbors.

Phương pháp Top-N (hình bên phải) không dựa vào khoảng cách, mà nó mở rộng đến khi có đủ N neighbors.

Như vậy có thể nói phương pháp Top-N ưu tiên số lượng hơn chất lượng, Threshold thì ngược lại.

Bước 5: Dự đoán điểm đánh giá. Công thức được tính như sau:

A picture containing text

Description automatically generated

Trong đó:

u : Điểm đánh giá trung bình của user u ( user hiên tại).

ru,j : Điểm đánh giá của user hiện tại dành cho item j.

Si : Các item trong vùng lân cận ( neighborhood) và đã được đánh giá bởi user hiện tại

Pred(u,i) : Điểm dự đoán của user hiện tại với item i.

sim(i,j) : Độ tương tự giữa hai item i và j.

### **Nhận xét hai cách tiếp cận User-based filtering và Item-based filtering**

Cách tiếp cần User-based filtering có một số hạn chế sau:

* Số lượng users thường lớn hơn số lượng items rất nhiều, dẫn đến kích thước ma trận tương đồng rất lớn .
* Người dùng thường ít sẵn lòng đánh giá, đồng nghĩa với việc ma trận đánh giá có nhiều phần tử mang giá trị rỗng, dẫn đến kết quả gợi ý kém hiệu quả hơn. Đồng thời mỗi khi user thêm hoặc sửa đánh giá, trung bình cộng các rating các rating cũng như vector chuẩn hóa ứng với user này thay đổi nhiều, dẫn đến phải thường xuyên cập nhật ma trận tương đồng, gây tốn nhiều bộ nhớ và thời gian.

Ngược lại, cách tiếp cận Item-based filtering có những lợi ích sau:

* Số lượng item thường nhỏ hơn số lượng user nên kích thước ma trận tương đồng nhỏ hơn nhiều.
* Vì mỗi item thường được đánh giá bởi nhiều user, nên giá trị trung bình của mỗi cột(items) ít bị ảnh hưởng, việc cập nhật ma trận tương đồng cũng có thể ít được cập nhật thường xuyên hơn.
  1. **Thuật toán TF-IDF.**

TF-IDF (Term Frequency – Inverse Document Frequency) là 1 kĩ thuật sử dụng trong khai phá dữ liệu văn bản. Trọng số này được sử dụng để đánh giá tầm quan trọng của một từ trong một văn bản. Giá trị cao thể hiện độ quan trọng cao và nó phụ thuộc vào số lần từ xuất hiện trong văn bản nhưng bù lại bởi tần suất của từ đó trong tập dữ liệu.

Một vài biến thể của tf-idf thường được sử dụng trong các hệ thống tìm kiếm như một công cụ chính để đánh giá và sắp xếp văn bản dựa vào truy vấn của người dùng.  Tf-idf cũng được sử dụng trong các bài toán như tóm tắt văn bản và phân loại văn bản.

**TF**: Term Frequency(Tần suất xuất hiện của từ) là số lần từ xuất hiện trong văn bản. Vì các văn bản có thể có độ dài ngắn khác nhau nên một số từ có thể xuất hiện nhiều lần trong một văn bản dài hơn là một văn bản ngắn. Như vậy, term frequency thường được chia cho độ dài văn bản( tổng số từ trong một văn bản).

A picture containing text, watch, gauge

Description automatically generated

Trong đó:

c(w,d): Tần số xuất hiện của từ vựng w trong văn bản d.

len(d): Số từ vựng trong văn bản d.

**IDF**: Inverse Document Frequency(Nghịch đảo tần suất của văn bản), giúp đánh giá tầm quan trọng của một từ . Khi tính toán TF , tất cả các từ được coi như có độ quan trọng bằng nhau. Nhưng  một số từ như “is”, “of” và “that” thường xuất hiện rất nhiều lần nhưng độ quan trọng là không cao. Như thế cần giảm độ quan trọng của những từ này xuống.

A picture containing logo

Description automatically generated

Trong đó:

M: Tổng số văn bản trong tập văn bản D.

f(w,D): Tần số của từ w trong tập văn bản D.

* 1. **Chức năng tìm kiếm các nhà hàng/quán ăn trên bản đồ**
  2. **Các công cụ hỗ trợ hệ thống**
     1. **Python**
     2. **Php**

# CHƯƠNG 3: PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

## **3.1. Phân tích hiện trạng nhu cầu**

### **3.1.1. Yêu cầu chức năng**

* Người dùng:
  + Xem thông tin nhà hàng/quán ăn.
  + Tạo nhà hàng/quán ăn.
  + Đánh giá và bình luận nhà hàng/quán ăn.
  + Điều khiển bằng giọng nói.
  + Chia sẻ nhà hàng/quán ăn..
* Hệ thống
  + Đề xuất địa điểm dựa trên lịch sử đánh giá của người dùng.
  + Xử lý điều khiển bằng giọng nói.
  + Gợi ý nhà hàng lên map.
* Người quản lý:
  + Quản lý nhà hàng/quán ăn.
  + Quản lý người dùng.
  + Tạo biểu mẫu thống kê.

### **3.1.2. Yêu cầu phi chức năng:**

* Giao diện thân thiện.
* Ứng dụng tiện ích.
* Tốc độ xử lý nhanh.

## **3.2. Phân tích và thiết kế thành phần dữ liệu**

### **3.2.1. Cấu trúc dữ liệu**

customers

Id

Avatar

DisplayName

IsFollow

IsVerified

Level

Rank

TotalPictures

TotalReviews

Url

UrlAlbums

UrlReviews

UserName

VerifyingPercent

Status

restaurants

Id

Address

AvgRating

Distance

HasPromotion

IsOpening

Latitude

Longitude

LocationUrlRewriteName

Name

ResCreatedOn

PhotoUrl

PriceMax

PriceMaxDisplay

PriceMin

PriceMinDisplay

PromotionId

PromotionTitle

PromotionURL

ResUrlAlbums

ResUrlReviews

Status

RestaurantStatus

TotalCheckIns

TotalFavourites

TotalPictures

TotalReviews

TotalSaves

TotalServices

Url

UrlRewriteName

resraurant\_details

id

res\_id

street\_address

district

city

price

category\_id

cuisine\_id

target\_customer\_id

open\_time

categories

id

name

status

parent\_id

comments

AvgRating

CreatedDate

CreatedOnTimeDiff

Description

DeviceName

DeviceType

DeviceUrl

Id

IsFirstUserReview

IsLiked

Owner\_id

Resid

Title

TotalLike

TotalPictures

TotalView

TypeName

Url

comment\_likes

Id

IdOwner

IdComment

Status

comment\_pictures

Id

BgColor

Description

Height

PhotoDetailUrl

TotalLikes

Url

Width

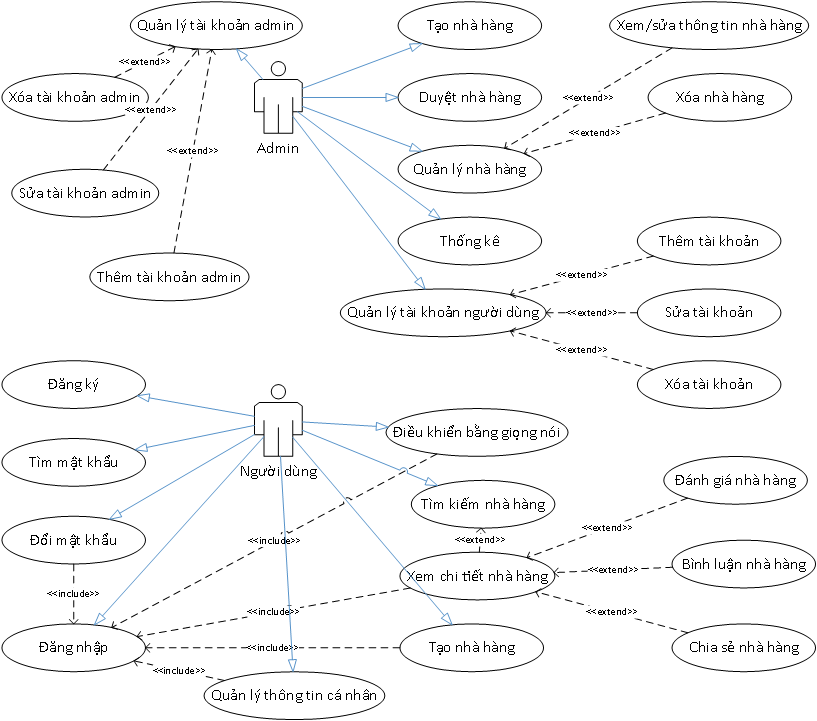
CommentId

### **3.2.2. Mô tả các thuộc tính**

**Chưa làm**

## **3.3. Phân tích thiết kế thành phần xử lý**

### **3.3.1. Sơ đồ Usecase**



### **3.3.2. Danh sách Usecase**

|  |  |  |
| --- | --- | --- |
| **STT** | **Usecase** | **Ý nghĩa** |
| 1 | Usecase Quản lý tài khoản admin(Admin) | Admin quản lý tài khoản của mình và các admin |
| 2 | Usecase Thêm tài khoản admin(Admin) | Admin thêm các tài khoản có quyền admin |
| 3 | Usecase Sửa tài khoản admin(Admin) | Admin xem/sửa thông tin các tài khoản admin |
| 4 | Usecase Xóa tài khoản admin(Admin) | Admin xóa các tài khoản admin |
| 5 | Usecase Quản lý tài khoản người dùng(Admin) | Admin quản lý tài khoản của người dùng |
| 6 | Usecase Thêm tài khoản(Admin) | Admin thêm tài khoản người dùng |
| 7 | Usecase Sửa tài khoản(Admin) | Admin xem/sửa thông tin tài khoản người dùng |
| 8 | Usecase Xóa tài khoản(Admin) | Admin xóa tài khoản người dùng |
| 9 | Usecase Tạo nhà hàng(Admin) | Admin tạo nhà hàng/quán ăn mới |
| 10 | Usecase Duyệt nhà hàng(Admin) | Admin duyệt các nhà hàng do người dùng tạo |
| 11 | Usecase Quản lý nhà hàng(Admin) | Admin quản lý việc xem/xóa/sửa nhà hàng/quán ăn |
| 12 | Usecase Xem/Sửa nhà hàng(Admin) | Admin xem/sửa thông tin nhà hàng/quán ăn |
| 13 | Usecase Xóa nhà hàng(Admin) | Admin xóa nhà hàng/quán ăn |
| 14 | Usecase Thống kê(Admin) | Admin thống kê thông tin hệ thống |
| 15 | Usecase Đăng ký(Người dùng) | Người dùng đăng ký tài khoản |
| 16 | Usecase Tìm mật khẩu(Người dùng) | Người dùng tìm mật khẩu của mình |
| 17 | Usecase Đổi mật khẩu(Người dùng) | Người dùng đổi mật khẩu của mình |
| 18 | Usecase Đăng nhập(Người dùng) | Người dùng đăng nhập vào website |
| 19 | Usecase Quản lý thông tin cá nhân(Người dùng) | Người dùng quản lý thông tin cá nhân |
| 20 | Usecase Điều khiển bằng giọng nói(Người dùng) | Người dùng điều khiển website bằng giọng nói |
| 21 | Usecase Tìm kiếm nhà hàng(Người dùng) | Người dùng tìm kiếm nhà hàng/quán ăn |
| 22 | Usecase Xem chi tiết nhà hàng(Người dùng) | Người dùng xem chi tiết nhà hàng/quán ăn |
| 23 | Usecase Đánh giá nhà hàng(Người dùng) | Người dùng đánh giá nhà hàng/quán ăn đang xem |
| 24 | Usecase Bình luận nhà hàng(Người dùng) | Người dùng bình luận nhà hàng/quán ăn đang xem |
| 25 | Usecase Chia sẻ nhà hàng(Người dùng) | Người dùng chia sẻ nhà hàng/quán ăn mình muốn |
| 26 | Usecase Tạo nhà hàng(Người dùng) | Người dùng tạo nhà hàng/quán ăn đề xuẩt cho website |

### **3.3.3. Sơ đồ tuần tự**

* Usecase Quản lý tài khoản admin:

Graphical user interface

Description automatically generated

* Usecase Thêm tài khoản admin

Diagram

Description automatically generated

* Usecase Sửa tài khoản admin

Diagram

Description automatically generated

* Usecase Xóa tài khoản admin

Graphical user interface

Description automatically generated

* Usecase Quản lý tài khoản

Graphical user interface

Description automatically generated

* Usecase Thêm tài khoản

Diagram

Description automatically generated

* Usecase Sửa tài khoản

Diagram

Description automatically generated

* Usecase Xóa tài khoản

Graphical user interface

Description automatically generated

* Usecase Tạo nhà hàng

Diagram

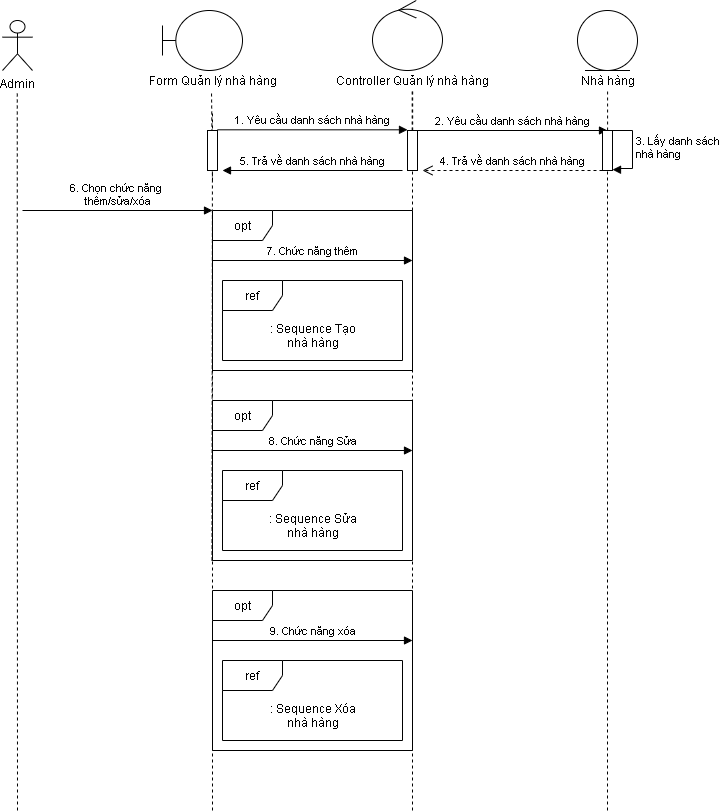
Description automatically generated

* Usecase Duyệt nhà hàng

Diagram

Description automatically generated

* Usecase Quản lý nhà hàng



* Usecase Xem/sửa nhà hàng

Diagram

Description automatically generated

* Usecase Xóa nhà hàng

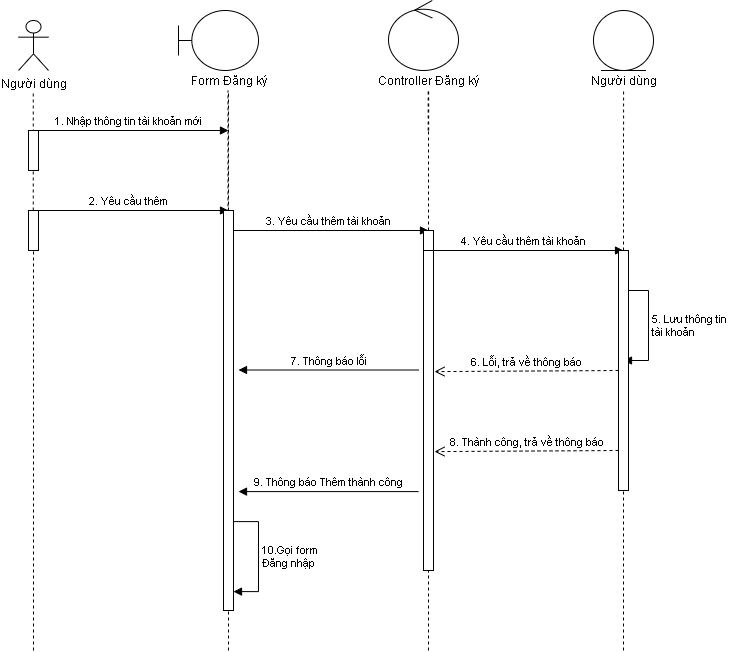
Graphical user interface

Description automatically generated

* Usecase Thống kê nhà hàng

**Chưa làm**

* Usecase Đăng ký



* Usecase Tìm mật khẩu

**Chưa làm**

* Usecase Đăng nhập

Diagram

Description automatically generated

* Usecase Đổi mật khẩu

Graphical user interface

Description automatically generated

* Usecase Quản lý thông tin cá nhân

**Chưa làm**

* Usecase Tạo nhà hàng(Người dùng)

Diagram

Description automatically generated

* Usecase Điều khiển bằng giọng nói

**Chưa làm**

* Usecase Tìm kiếm nhà hàng

Diagram

Description automatically generated

* Usecase Xem chi tiết nhà hàng

A picture containing diagram

Description automatically generated

* Usecase Đánh giá

**Chưa làm**

* Usecase Bình luận

**Chưa làm**

* Usecase Chia sẻ

# CHƯƠNG 4: CÀI ĐẶT VÀ TRIỂN KHAI HỆ THỐNG

## 4.1. **Môi trường triển khai**

## **4.2. Giải pháp công nghệ phía Client**

## **4.3. Giải pháp công nghệ phía Server**

### **4.3.1. Ứng dụng thuật toán lọc cộng tác:**

Xây dựng ma trận đánh giá:

# CHƯƠNG 5: KẾT LUẬN

**Chưa làm**

# TÀI LIỆU THAM KHẢO

**Chưa làm**