

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN**

LÊ THỊ PHƯƠNG LINH – BÙI ĐĂNG KHOA

**HỆ THỐNG GỢI Ý SẢN PHẨM DỰA
TRÊN DỊCH VỤ TƯƠNG TÁC CHO CÁC
DOANH NGHIỆP VỪA VÀ NHỎ**

**KHÓA LUẬN TỐT NGHIỆP CỬ NHÂN CNTT
CHƯƠNG TRÌNH CHẤT LƯỢNG CAO**

THÀNH PHỐ HỒ CHÍ MINH, 2023

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN**

LÊ THỊ PHƯƠNG LINH - MSSV: 19127041

BÙI ĐĂNG KHOA - MSSV: 19127645

**HỆ THỐNG GỢI Ý SẢN PHẨM DỰA
TRÊN DỊCH VỤ TƯƠNG TÁC CHO CÁC
DOANH NGHIỆP VỪA VÀ NHỎ**

**KHÓA LUẬN TỐT NGHIỆP CỬ NHÂN CNTT
CHƯƠNG TRÌNH CHẤT LƯỢNG CAO**

GIẢNG VIÊN HƯỚNG DẪN

TS. PHẠM NGUYỄN CƯỜNG

TS. VŨ THỊ MỸ HẰNG

THÀNH PHỐ HỒ CHÍ MINH, 2023

LỜI CẢM ƠN

Trải qua thời gian học tập tại trường Đại học Khoa học Tự nhiên – ĐHQG-HCM, chúng tôi lựa chọn làm luận văn tốt nghiệp với đề tài “Hệ thống gợi ý sản phẩm dựa trên dịch vụ tương tác cho các doanh nghiệp vừa và nhỏ” để tổng hợp lại kiến thức của mình. Trong suốt quá trình làm khóa luận, chúng tôi đã nhận được sự hướng dẫn, giúp đỡ quý báu từ nhà trường, thầy cô, gia đình và bạn bè.

Lời đầu tiên, chúng tôi xin chân thành cảm ơn thầy Phạm Nguyễn Cương và cô Vũ Thị Mỹ Hằng đã tận tình hướng dẫn, chỉ bảo và luôn có sự phản hồi tỉ mỉ trong thời gian nhanh nhất nhằm giúp chúng tôi có thể thực hiện khóa luận một cách tốt nhất.

Chúng tôi cũng xin cảm ơn Ban chủ nhiệm Khoa và tất cả các thầy cô của Khoa Công nghệ thông tin đã tạo điều kiện tốt nhất giúp chúng tôi hoàn thành khóa luận.

Ngoài ra, xin cảm ơn tất cả thầy cô trường Đại học Khoa học Tự nhiên, ĐHQG-HCM, cũng như bạn bè, gia đình đã động viên, cổ vũ để chúng tôi để chúng tôi vượt qua những khó khăn trong quá trình thực hiện khóa luận.

Do hạn chế về thời gian nên sẽ không tránh khỏi những thiếu sót trong nghiên cứu, vì vậy, chúng tôi xin lắng nghe và tiếp thu những ý kiến phản biện của thầy cô để khắc phục sai sót, bổ sung kiến thức cho bản thân.

Cuối cùng, chúng tôi xin kính chúc thầy Phạm Nguyễn Cương, cô Vũ Thị Mỹ Hằng, các thầy cô Ban lãnh đạo và các phòng ban chức năng Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM dồi dào sức khỏe và thành công trong sự nghiệp.

Chúng tôi xin chân thành cảm ơn!

Lê Thị Phương Linh

Bùi Đăng Khoa

LỜI CAM KẾT

Chúng tôi xin cam đoan rằng đây là công trình nghiên cứu của chúng tôi, có sự hỗ trợ từ Giáo viên hướng dẫn là TS. Phạm Nguyễn Cương và TS. Vũ Thị Mỹ Hằng.

Những số liệu, bảng biểu phục vụ cho việc phân tích và dẫn dắt đề tài này được thu thập từ các nguồn tài liệu khác nhau được ghi chú trong mục tài liệu tham khảo hoặc chú thích ngay bên dưới các bảng biểu.

Ngoài ra, đối với các tài liệu diễn giải để làm rõ thêm các luận điểm đã phân tích và trích dẫn trong phần phụ lục cũng được chú thích nguồn gốc dữ liệu.

Nếu phát hiện có bất kỳ sự gian lận nào chúng tôi xin hoàn toàn chịu trách nhiệm trước Hội đồng cũng như kết quả luận văn của mình.

ĐỀ CƯƠNG CHI TIẾT

Tên đề tài:

HỆ THỐNG GỢI Ý SẢN PHẨM DỰA TRÊN DỊCH VỤ TƯƠNG TÁC
CHO DOANH NGHIỆP VỪA VÀ NHỎ

Giảng viên hướng dẫn: TS. Phạm Nguyễn Cương – TS. Vũ Thị Mỹ Hằng

Thời gian thực hiện: 9/2022 – 4/2023

Sinh viên thực hiện: 19127041 – Lê Thị Phương Linh, 19127645 – Bùi Đăng Khoa

Loại đề tài: Nghiên cứu

Nội dung luận văn

Mục tiêu:

- **Lý thuyết:**

- Tìm hiểu khái niệm về phần mềm hướng dịch vụ (SaaS), dịch vụ tương tác.
- Khảo sát các phần mềm hướng dịch vụ phổ biến hiện nay.
- Tìm hiểu mô hình của hệ thống gợi ý sản phẩm.
- Tìm hiểu các thuật toán gợi ý sản phẩm phổ biến: Lọc cộng tác (Collaborative Filtering), dựa trên nhân khẩu học (Demographic), dựa trên nội dung (Content-based), gợi ý sản phẩm phổ biến (Hot items).

- **Ứng dụng:** Triển khai hệ thống gợi ý sản phẩm vừa và nhỏ dựa trên dịch vụ tương tác, hỗ trợ các doanh nghiệp vừa và nhỏ tăng trải nghiệm khách hàng, đồng thời tăng năng suất kinh doanh. Đây là giải pháp mới giúp vượt qua những hạn chế của các hệ thống gợi ý sản phẩm hiện tại.

Phương pháp thực hiện:

- **Lý thuyết:**

- Tìm hiểu các tài liệu khoa học đã công bố về mô hình hệ thống gợi ý sản phẩm.

- Kiểm tra các công cụ, nền tảng hệ thống gợi ý sản phẩm hiện tại để nắm các chức năng chính cần cung cấp cũng như hạn chế cần khắc phục.
- Đề xuất giải pháp, phân tích, thiết kế, triển khai hệ thống gợi ý sản phẩm theo hướng dịch vụ kết hợp đa dạng thuật toán gợi ý sản phẩm: Lọc cộng tác (Collaborative Filtering), dựa trên nhân khẩu học (Demographic), dựa trên nội dung (Content-based), gợi ý sản phẩm phổ biến (Hot items).

- **Kỹ thuật:**

- Sử dụng Django để phát triển máy chủ gợi ý sản phẩm.
- Sử dụng ReactJS để xây dựng giao diện quản lý cho admin website thương mại điện tử.
- Sử dụng các cơ sở dữ liệu quan hệ phổ biến như PostgreSQL, MySQL, Oracle, MicrosoftSQL cho hệ thống quản lý dữ liệu.
- Sử dụng Airflow để cài đặt tiến trình thực hiện các tác vụ trong hệ thống.
- Sử dụng Digital Ocean để lưu trữ dữ liệu trên Cloud.
- Sử dụng thư viện Tensorflow, Gensim, Scikit-learn để triển khai các thuật toán gợi ý sản phẩm.
- Sử dụng RESTful API để xử lý các yêu cầu từ người dùng.
- Ngôn ngữ lập trình: Python, Javascript.

Kết quả đạt được:

- Tiếp thu kiến thức về phần mềm hướng dịch vụ và dịch vụ tương tác và ảnh hưởng của chúng trong lĩnh vực thương mại điện tử.
- Triển khai hệ thống gợi ý sản phẩm theo hướng dịch vụ và tích hợp vào các website thương mại điện tử.
- Áp dụng các thuật toán gợi ý sản phẩm vào hệ thống.
- Sử dụng các nền tảng và công cụ phổ biến vào việc xây dựng hệ thống.





Kế hoạch thực hiện:

| STT | Công việc | Thời gian |
|------------|--|-------------------------|
| 1 | Nghiên cứu hệ thống gợi ý | 01/09/2022 - 30/09/2022 |
| 2 | Nghiên cứu các mô hình máy học gợi ý | 01/09/2022 - 30/09/2022 |
| 3 | Nghiên cứu phần mềm hướng dịch vụ | 01/09/2022 – 30/09/2022 |
| 4 | Khảo sát các trang thương mại điện tử và các doanh nghiệp vừa và nhỏ | 01/10/2022 - 30/10/2022 |
| 5 | Thiết kế ERD cho hệ thống | 01/10/2022 - 30/10/2022 |
| 6 | Thiết kế kiến trúc cho hệ thống | 01/10/2022 - 30/10/2022 |
| 7 | Xây dựng mô hình máy học dựa trên các thư viện có sẵn | 01/11/2022 - 30/11/2022 |
| 8 | Tạo dữ liệu mẫu | 01/11/2022 - 30/11/2022 |
| 9 | Xây dựng website TMĐT (ReactJS) | 01/12/2022 – 31/12/2022 |
| 10 | Xây dựng website TMĐT (Wordpress) | 01/12/2022 – 31/12/2022 |
| 11 | Xây dựng chức năng quản lý | 01/01/2023 – 30/01/2023 |
| 12 | Xây dựng chức năng bắt sự kiện hành vi trên website TMĐT | 01/01/2023 – 30/01/2023 |
| 13 | Xây dựng dịch vụ gợi ý sản phẩm | 01/02/2023 – 28/02/2023 |

| STT | Công việc | Thời gian |
|-----|---|-------------------------|
| 14 | Deploy hệ thống trên Digital Ocean | 01/02/2023 – 28/02/2023 |
| 15 | Kiểm chứng trên các trang TMĐT tự xây dựng và sửa lỗi | 01/02/2023 – 28/02/2023 |
| 16 | Viết báo cáo | 01/02/2023 – 05/03/2023 |

| | |
|-----------------------|---|
| Xác nhận của GVHD | TP Hồ Chí Minh, ngày 06 tháng 03 năm 2023 |
| <i>Chữ ký GVHD</i> | <i>Chữ ký nhóm SV thực hiện</i> |
| TS. Phạm Nguyễn Cương | Lê Thị Phương Linh |
| TS. Vũ Thị Mỹ Hằng | Bùi Đăng Khoa |

| STT | Công việc | Thời gian |
|-----|---|-------------------------|
| 14 | Deploy hệ thống trên Digital Ocean | 01/02/2023 – 28/02/2023 |
| 15 | Kiểm chứng trên các trang TMĐT tự xây dựng và sửa lỗi | 01/02/2023 – 28/02/2023 |
| 16 | Viết báo cáo | 01/02/2023 – 05/03/2023 |

| | | | |
|---|--|--|--|
| <p>Xác nhận của GVHD</p> <p>Chữ ký GVHD</p>  <p>TS. Phạm Nguyễn Cương</p>  <p>TS. Vũ Thị Mỹ Hằng</p> | | <p>TP Hồ Chí Minh, ngày 06 tháng 03 năm 2023</p> <p>Chữ ký nhóm SV thực hiện</p>  <p>Lê Thị Phương Linh</p>  <p>Bùi Đăng Khoa</p> | |
|---|--|--|--|

MỤC LỤC

| | |
|---|------|
| LỜI CẢM ƠN | ii |
| LỜI CAM KẾT | iii |
| ĐỀ CƯƠNG CHI TIẾT | iv |
| MỤC LỤC | ix |
| BẢNG THUẬT NGỮ VIẾT TẮT | xii |
| DANH MỤC HÌNH ẢNH | xiii |
| DANH MỤC BẢNG BIỂU | xv |
| CẤU TRÚC LUẬN VĂN | xvi |
| CHƯƠNG 1: MỞ ĐẦU | 2 |
| 1.1. Giới thiệu đề tài | 2 |
| 1.2. Lý do chọn đề tài và mục tiêu đề tài | 4 |
| 1.3. Phạm vi đề tài | 6 |
| 1.4. Kết quả đạt được | 7 |
| CHƯƠNG 2: NỀN TẢNG LÝ THUYẾT VÀ CÔNG NGHỆ | 8 |
| 2.1. Phần mềm hướng dịch vụ (SaaS – Software-as-a-Service) | 8 |
| 2.1.1. Tổng quan | 8 |
| 2.1.2. Đánh giá chung | 11 |
| 2.2. Phần mềm gợi ý hướng dịch vụ (RaaS – Recommender-as-a-Service) | 11 |
| 2.2.1. Dịch vụ tương tác (Interactive service) | 11 |
| 2.2.2. Phần mềm gợi ý hướng dịch vụ | 16 |
| 2.3. Nghiên cứu thuật toán gợi ý | 18 |
| 2.3.1. Lọc cộng tác (Collaborative Filtering) | 20 |
| 2.3.2. Dựa trên nhân khẩu học (Demographic Filtering) | 21 |
| 2.3.3. Dựa trên nội dung (Content-based) | 22 |
| 2.3.4. Gợi ý sản phẩm phổ biến nhất (Hot items) | 23 |
| 2.3.5. Kết luận | 23 |
| 2.4. Công cụ và môi trường phát triển | 24 |

| | | |
|--|--|----|
| 2.4.1. | Hệ thống phụ trợ – Django | 24 |
| 2.4.2. | Giao diện khách hàng doanh nghiệp – ReactJS | 25 |
| 2.4.3. | Các cơ sở dữ liệu quan hệ | 25 |
| 2.4.4. | Bộ lập lịch – Airflow | 25 |
| 2.4.5. | Điện toán đám mây – Digital Ocean | 26 |
| 2.5. | Kết luận chương | 26 |
| CHƯƠNG 3: MÔ HÌNH TRIỂN KHAI DỊCH VỤ GỢI Ý | | 27 |
| 3.1. | Phân tích yêu cầu | 27 |
| 3.1.1. | Khảo sát hệ thống thực tế | 27 |
| 3.1.2. | Phân tích, kết luận | 32 |
| 3.2. | Framework Trivi | 33 |
| 3.2.1. | Nguyên tắc của hệ thống thông minh về khách hàng | 33 |
| 3.2.2. | Mô hình dữ liệu | 34 |
| 3.3. | Thiết kế giải pháp | 34 |
| 3.3.1. | Sơ đồ use cases. | 35 |
| 3.3.2. | Sơ đồ thực thể quan hệ (ERD) | 39 |
| 3.3.3. | Thiết kế kiến trúc hệ thống | 42 |
| 3.3.4. | Sơ đồ Sequence | 44 |
| 3.3.5. | Đánh giá kiến trúc hệ thống | 46 |
| 3.3. | Cài đặt giải pháp | 47 |
| 3.3.1. | Phương thức phân phối dịch vụ | 48 |
| 3.3.2. | Máy chủ gợi ý (Recommender server) | 50 |
| 3.3.3. | Máy chủ lập lịch (Scheduler server) | 52 |
| 3.3.4. | Bắt sự kiện hành vi người dùng | 52 |
| 3.4. | Kết luận chương | 55 |
| CHƯƠNG 4: XÂY DỰNG HỆ THỐNG VÀ THỰC NGHIỆM | | 56 |

| | |
|---|----|
| 4.1. Cài đặt package sử dụng | 56 |
| 4.2. Thiết lập và cấu hình | 57 |
| 4.3. Cấu trúc thư mục mã nguồn hệ thống | 60 |
| 4.3.1. Thành phần gợi ý sản phẩm | 60 |
| 4.3.2. Thành phần giao diện khách hàng doanh nghiệp | 62 |
| 4.3.3. Triển khai thành phần trên các máy chủ đám mây | 63 |
| 4.4. Thực nghiệm | 65 |
| 4.3.1. Kiểm chứng thuật toán trên website TMĐT | 66 |
| 4.3.2. Minh họa phân hệ quản lý | 70 |
| 4.5. Kết luận chương | 75 |
| CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN | 76 |
| 5.1. Kết quả đạt được | 76 |
| 5.2. Hạn chế của luận văn | 77 |
| 5.3. Hướng phát triển tương lai | 78 |
| 5.3.1. Phát triển kiến trúc | 78 |
| 5.3.2. Phát triển về hiệu quả thuật toán | 79 |
| 5.3.3. Phát triển về cách bắt sự kiện | 79 |
| 5.3.4. Kết luận | 80 |
| PHỤ LỤC | 81 |
| Mô tả thuộc tính của cơ sở dữ liệu | 81 |
| TÀI LIỆU THAM KHẢO | 87 |

BẢNG THUẬT NGỮ VIẾT TẮT

| ID | Ký hiệu chữ viết tắt | Chữ viết đầy đủ |
|----|----------------------|--------------------|
| 1 | CSDL | Cơ sở dữ liệu |
| 2 | TMĐT | Thương mại điện tử |

DANH MỤC HÌNH ẢNH

| | |
|--|----|
| Hình 1. Quy mô thị trường TMĐT bán lẻ tại Việt Nam [3] | 2 |
| Hình 2. Kiến trúc SaaS | 10 |
| Hình 3. Mô hình dịch vụ tương tác..... | 12 |
| Hình 4. Quy trình gợi ý hướng dịch vụ..... | 17 |
| Hình 5. Phân loại thuật toán gợi ý sản phẩm [14] | 19 |
| Hình 6. Sơ đồ Use Case Dịch vụ tư vấn sản phẩm | 35 |
| Hình 7. Lược đồ thực thể kết hợp – phần quản lý dữ liệu | 39 |
| Hình 8. Sơ đồ thực thể quan hệ (Entity Relationship Diagram) – phần quản lý tài khoản..... | 41 |
| Hình 9. Kiến trúc hệ thống..... | 44 |
| Hình 10. Sơ đồ tuần tự của hệ thống gợi ý | 44 |
| Hình 11. Sơ đồ tuần tự bộ lập lịch của hệ thống..... | 45 |
| Hình 12. Kiến trúc hệ thống mở rộng | 47 |
| Hình 13. Recommender mẫu thiết kế hướng đối tượng..... | 51 |
| Hình 14. Màn hình làm việc của Airflow | 52 |
| Hình 15. Cách copy XPath bằng trình duyệt | 53 |
| Hình 16. Trang chủ website Concung | 54 |
| Hình 17. Cách thức bắt sự kiện trên website TMĐT | 55 |
| Hình 18. Quy trình tạo dự án Django | 58 |
| Hình 19. Quy trình tạo dự án ReactJS | 58 |
| Hình 20. Cấu hình CSDL..... | 59 |
| Hình 21. Minh họa tạo bảng trong dự án Django | 60 |
| Hình 22. Cấu trúc quản lý phía hệ thống gợi ý sản phẩm..... | 60 |
| Hình 23. Cấu trúc quản lý phía giao diện khách hàng doanh nghiệp | 62 |
| Hình 24. Màn hình làm việc của Digital Ocean | 64 |
| Hình 25. Màn hình làm việc của Neon | 65 |
| Hình 26. Đoạn mã được nhúng vào website TMĐT (sử dụng ReactJS) | 66 |
| Hình 27. Trang chủ website TMĐT mẫu (ReactJS) | 67 |
| Hình 28. Cấu trúc JSON kết quả gợi ý sản phẩm | 68 |

| | |
|--|----|
| Hình 30. Kết quả minh họa trên website TMĐT (ReactJS)..... | 69 |
| Hình 29. Kết quả minh họa trên website TMĐT (Wordpress) | 69 |
| Hình 31. Minh họa pop-up hiển thị kết quả gợi ý sản phẩm | 70 |
| Hình 32. Giao diện đăng nhập (bên trái), đăng ký tài khoản (bên phải)..... | 71 |
| Hình 33. Giao diện Dashboard (khi chưa có dữ liệu) | 71 |
| Hình 34. Giao diện Import data. | 72 |
| Hình 35. Giao diện Dashboard | 72 |
| Hình 36. Giao diện Configuration | 73 |
| Hình 37. Cách lấy XPath | 74 |
| Hình 38. Giao diện xem dữ liệu hành vi Click | 74 |
| Hình 39. Minh họa mô hình Web Page Segmentation [21] | 80 |

DANH MỤC BẢNG BIỂU

| | |
|--|----|
| Bảng 1. Mô tả tác nhân trong hệ thống..... | 37 |
| Bảng 2. Đặc tả chức năng | 38 |
| Bảng 3. Mô tả thực thể..... | 41 |
| Bảng 4. Các packages cần thiết trong hệ thống | 57 |
| Bảng 5. Mô tả các thư mục chính phía hệ thống gợi ý sản phẩm | 61 |
| Bảng 6. Mô tả thư mục chính phía giao diện khách hàng doanh nghiệp | 63 |
| Bảng 7. Thông tin server thuê từ Digital Ocean | 64 |
| Bảng 8. Các CSDL của hệ thống | 65 |
| Bảng 9. Mô tả cấu trúc JSON kết quả gợi ý sản phẩm | 68 |
| Bảng 10. Mô tả bảng CUSTOMER..... | 81 |
| Bảng 11. Mô tả bảng CUSTOMER_PROFILE..... | 82 |
| Bảng 12. Mô tả bảng PRODUCT | 82 |
| Bảng 13. Mô tả bảng SESSION..... | 83 |
| Bảng 14. Mô tả bảng WEB_EVENT | 83 |
| Bảng 15. Mô tả bảng BUSINESS_TRANSACTION..... | 84 |
| Bảng 16. Mô tả bảng EVENT_ITEM..... | 84 |
| Bảng 17. Mô tả bảng TRANSACTION_ITEM..... | 85 |
| Bảng 18. Mô tả bảng SCHEDULER | 85 |
| Bảng 19. Mô tả bảng RECOMMENDER_MODEL | 85 |
| Bảng 20. Mô tả bảng RECOMMENDER_STRATEGY | 86 |

CẤU TRÚC LUẬN VĂN

Luận văn được tổ chức thành 5 chương:

CHƯƠNG 1: MỞ ĐẦU

Chương 1 nhằm phân tích bối cảnh thị trường thương mại điện tử, từ đó dẫn đến động lực phát triển đề tài của chúng tôi. Trong chương này, chúng tôi giới thiệu tổng quan về đề tài, lý do chọn đề tài, mục tiêu và phạm vi của đề tài. Tiếp theo, chúng tôi đưa ra các công việc chính để hoàn thiện luận văn này.

CHƯƠNG 2: NỀN TẢNG LÝ THUYẾT VÀ CÔNG NGHỆ

Chương 2 trình bày tổng quan về nền tảng lý thuyết và công nghệ được sử dụng trong luận văn. Trong chương này, đầu tiên chúng tôi trình bày định nghĩa và đánh giá vai trò của phần mềm hướng dịch vụ (SaaS – Software-as-a-Software) trong cuộc sống hiện nay. Từ đó, chúng tôi trình bày ý tưởng triển khai phần mềm gợi ý hướng dịch vụ và phân tích những ảnh hưởng mà nó mang lại. Sau cùng, chúng tôi trình bày các khái niệm về thuật toán gợi ý và giới thiệu công cụ, môi trường lập trình. Chương này được xem như là bước chuẩn bị kiến thức để thực hiện luận văn.

CHƯƠNG 3: MÔ HÌNH TRIỂN KHAI DỊCH VỤ GỢI Ý

Chương 3 trình bày chi tiết về mô hình gợi ý sản phẩm hướng dịch vụ mà chúng tôi đã đề xuất trong luận văn. Từ kiến thức đã nghiên cứu ở chương 2, trong chương 3, trước hết, chúng tôi sẽ khảo sát hiện trạng các dịch vụ gợi ý hiện tại, từ đó phân tích để xác định các yêu cầu của một hệ thống gợi ý sản phẩm hướng dịch vụ. Giải pháp của chúng tôi được triển khai dựa trên Framework Trivi, đã được công bố trong bài báo Trivi: A conceptual Framework for Customer Intelligence Systems for Small and Medium-sized Enterprises [6]. Nguyên tắc cơ bản và mô hình dữ liệu được đề xuất từ hệ thống Trivi cũng sẽ được giới thiệu sơ lược trong chương này. Sau cùng, chúng tôi sẽ trình bày giải pháp chúng tôi đề

xuất cho mô hình triển khai dịch vụ gợi ý dựa trên nền tảng Framework Trivi, bao gồm kiến trúc hệ thống, mô hình cơ sở dữ liệu và sơ đồ tương tác (sequence diagram) của các thành phần trong hệ thống.

CHƯƠNG 4: XÂY DỰNG HỆ THỐNG VÀ THỰC NGHIỆM

Chương 4 nhằm minh họa các bước xây dựng hệ thống và tiến hành thực nghiệm để minh họa kết quả. Trong chương này, vận dụng các kiến thức và mô hình đã trình bày ở các chương trước, chúng tôi thiết lập và cấu hình hệ thống, tổ chức các thành phần hệ thống và lập trình các thành phần phục vụ luận văn. Sau cùng, chúng tôi tiến hành thực nghiệm trên website TMĐT mẫu để kiểm chứng kết quả.

CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

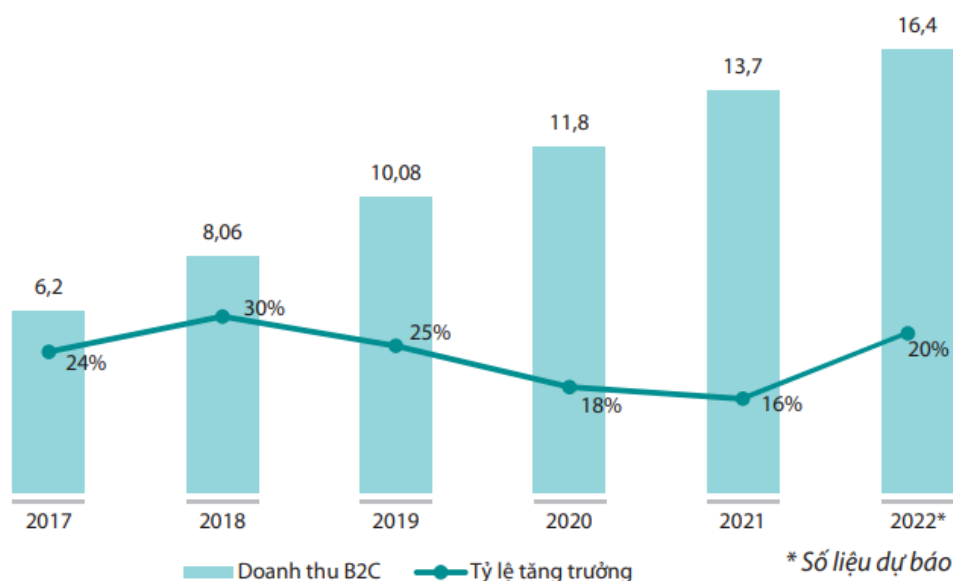
Chương 5 trình bày kết luận, đóng góp của luận văn, cũng như các đánh giá về quá trình hoàn thiện luận văn. Chúng tôi trước hết trình bày cụ thể kết quả đạt được của luận văn về mặt lý thuyết cũng như thực nghiệm. Các kiến thức đạt được cũng như những hạn chế chưa khắc phục được. Cuối cùng chúng tôi trình bày hướng phát triển của đề tài trong tương lai.

CHƯƠNG 1: MỞ ĐẦU

Chương 1 nhằm phân tích bối cảnh thị trường thương mại điện tử, từ đó dẫn đến động lực phát triển đề tài của chúng tôi. Trong chương này, chúng tôi giới thiệu tổng quan về đề tài, lý do chọn đề tài, mục tiêu và phạm vi của đề tài. Tiếp theo, chúng tôi đưa ra các công việc chính để hoàn thiện luận văn này.

1.1. Giới thiệu đề tài

Thương mại điện tử (TMĐT) là hình thức kinh doanh trực tuyến sử dụng nền tảng công nghệ thông tin để thực hiện các giao dịch mua bán, trao đổi, thanh toán trực tuyến. Với tốc độ phát triển nhanh chóng của mạng máy tính, TMĐT dần trở thành xu hướng kỷ nguyên số. Hơn thế nữa, sự xuất hiện của đại dịch Covid-19 gần đây đã dịch chuyển thói quen mua sắm của khách hàng từ mua bán trực tiếp sang online, góp phần khẳng định sự phổ biến và tầm quan trọng của TMĐT trong cuộc sống hằng ngày [1][2].



Hình 1. Quy mô thị trường TMĐT bán lẻ tại Việt Nam [3]

Trong bối cảnh COVID-19 vừa qua, một số ngành dịch vụ ghi nhận tăng trưởng âm đã làm giảm mức tăng chung của khu vực dịch vụ và toàn bộ nền kinh tế. Tuy nhiên, theo xu hướng phát triển của TMĐT toàn cầu, quy mô thị trường TMĐT tại Việt Nam ghi nhận tăng trưởng bền vững. Theo thống kê của Bộ Công Thương Việt Nam, thị trường TMĐT bán lẻ đều tăng hằng năm trên mức 16% từ năm 2017 đến 2022 [3]. Nhóm nghiên cứu Statista đã đánh giá rằng ngành TMĐT nước ta vào năm 2025 có thể tăng gấp 3 lần so với năm 2021, đạt 39 tỉ USD [4].

Với sự phát triển của công nghệ trong ngành TMĐT ngày nay, khách hàng có thể dễ dàng mua sắm trên các trang TMĐT với số lượng mặt hàng đa dạng, tuy nhiên, lượng dữ liệu khổng lồ ấy gây quá tải cho khách hàng khi thao tác trực tuyến. Để tối đa hóa trải nghiệm của khách hàng, việc gợi ý sản phẩm phù hợp là rất cần thiết. Tương tự như những lời tư vấn trực tiếp của hình thức mua hàng truyền thống, một hệ thống gợi ý tốt sẽ giúp khách hàng dễ dàng mua sản phẩm phù hợp, từ đó, tăng hiệu suất của hình thức mua hàng trực tuyến. Các hệ thống gợi ý sản phẩm được xây dựng để giảm bớt tình trạng quá tải và gợi ý các sản phẩm phù hợp nhất cho khách hàng, đáp ứng đúng nhu cầu của khách hàng [5].

Một hệ thống gợi ý mạnh mẽ góp phần giúp doanh nghiệp cải thiện doanh số bán hàng, giữ chân và thấu hiểu khách hàng. Những lợi ích đó đều là nhu cầu của tất cả doanh nghiệp. Tuy nhiên theo khảo sát của chúng tôi, xây dựng hệ thống gợi ý mạnh mẽ đang là cuộc chơi của những “ông lớn” điển hình như Amazon, Netflix, Shopee, v.v. Với nguồn lực và kinh phí hạn hẹp, các doanh nghiệp vừa và nhỏ khó có thể tham gia vào cuộc chơi này.

Kế thừa ý tưởng xây dựng hệ thống gợi ý (recommender system) từ bài báo Trivi: A conceptual framework for customer intelligence systems for small and medium-sized enterprises [6], trong luận văn này, chúng tôi nghiên cứu cung cấp dịch vụ gợi ý sản phẩm được tích hợp vào các trang thương mại điện tử của doanh nghiệp, hướng đến dịch vụ tương tác (interactive service) nhằm cải thiện

trải nghiệm khách hàng. Chúng tôi sẽ trình bày tổng quan về kỹ thuật gợi ý với kiểm chứng cụ thể trên các trang TMĐT mà chúng tôi xây dựng.

1.2. Lý do chọn đề tài và mục tiêu đề tài

Hệ thống gợi ý sản phẩm sẽ dựa trên thông tin về sản phẩm và hành vi của khách hàng để dự đoán các sản phẩm mà khách hàng có thể quan tâm từ đó giúp gia tăng trải nghiệm của khách hàng cũng như doanh thu và khả năng cạnh tranh của doanh nghiệp. Tuy nhiên, hiện nay các doanh nghiệp ở Việt Nam, đặc biệt là các doanh nghiệp vừa và nhỏ, vẫn còn nhiều khó khăn và hạn chế trong việc vận hành hệ thống gợi ý do điều kiện về tài nguyên và nguồn lực có nhiều giới hạn.

Hệ thống gợi ý sản phẩm đã có từ lâu và các thuật toán gợi ý (hay tư vấn) cũng được nghiên cứu rất nhiều [7][8][9] nhưng vẫn chưa có một mô hình giải pháp tổng quát được đưa ra để hỗ trợ các doanh nghiệp TMĐT. Mô hình giải pháp hướng dịch vụ trong đó các chức năng tư vấn được cung cấp dưới dạng dịch vụ ở nhiều mức độ khác nhau (phần mềm hướng dịch vụ, kiến trúc hướng dịch vụ, ...) [10][11] có thể được xem là một giải pháp khả thi.

Với mong muốn trên, mục tiêu đề tài của chúng tôi hướng tới việc đề xuất một mô hình hướng dịch vụ cho hệ thống gợi ý sản phẩm có thể tích hợp vào các trang thương mại điện tử của các doanh nghiệp vừa và nhỏ. Theo hướng tiếp cận này, chúng tôi áp dụng các thuật toán gợi ý sản phẩm đã được công bố và các chức năng chính của một hệ thống gợi ý sản phẩm (đăng ký dịch vụ, gợi ý sản phẩm, quản lý dữ liệu, ...) sẽ được chúng tôi đóng gói và cung cấp cho các doanh nghiệp vừa và nhỏ dưới dạng phần mềm dịch vụ. Dữ liệu sẽ được lưu trữ trên đám mây để tăng năng suất, tiết kiệm thời gian và chi phí. Mô hình này giúp cho các doanh nghiệp vừa và nhỏ, với nhiều hạn chế về tài nguyên và nguồn lực trong việc tự triển khai một hệ thống gợi ý sản phẩm riêng biệt, có thể tích hợp được

dịch vụ gợi ý sản phẩm trong các trang TMĐT của họ để gia tăng trải nghiệm khách hàng.

Để đạt được mục tiêu trên, công việc chính sẽ được triển khai trong phạm vi đề tài này được liệt kê như bên dưới:

- Nghiên cứu nền tảng lý thuyết

Để xây dựng một hệ thống gợi ý sản phẩm, chúng tôi tiếp cận và nghiên cứu những thuật toán đã được công nhận: lọc cộng tác (collaborative filtering), nhân khẩu học (demographic), nội dung (content-based). Bên cạnh nhóm thuật toán theo hướng cá nhân hóa, chúng tôi đề xuất triển khai thuật toán gợi ý sản phẩm theo độ phổ biến dựa trên tương tác của khách hàng. Ngoài ra, chúng tôi tìm hiểu về mô hình hướng dịch vụ nhằm cung cấp đầy đủ kiến thức phục vụ cho đề tài này.

- Tìm hiểu nền tảng công nghệ

Với vô số nền tảng công nghệ đang có trên thị trường, chúng tôi đã chọn lọc một vài công cụ và môi trường thích hợp với hệ thống gợi ý sản phẩm để triển khai. Các nền tảng chúng tôi sử dụng rất phổ biến hiện nay như React, Django, Airflow, và các cơ sở dữ liệu quan hệ như PostgreSQL, MySQL, MicrosoftSQL. Các công cụ này góp phần giúp chúng tôi thực hiện luận văn nhanh chóng, ổn định và hiệu quả.

- Đề xuất mô hình hệ thống gợi ý sản phẩm hướng dịch vụ

Từ việc khảo sát một số hệ thống gợi ý hiện có, kết hợp với kiến thức đã tìm hiểu, chúng tôi đề xuất mô hình gợi ý sản phẩm hướng dịch vụ với kiến trúc và phương thức phân phối dịch vụ sẽ được phân tích rõ ở chương 3.

- Thực nghiệm để minh họa cho mô hình

Chúng tôi xây dựng các website TMĐT để thử nghiệm với dữ liệu mẫu tự cung cấp, deploy hệ thống gợi ý sản phẩm và tiến hành nhúng đoạn mã vào website để minh họa cho mô hình đề xuất.

1.3. Phạm vi đề tài

- Đối tượng nghiên cứu của luận văn:

Các trang thương mại điện tử của doanh nghiệp vừa và nhỏ (Số lượng sản phẩm trên trang nhỏ hơn 2000).

Trong phạm vi luận văn này, chúng tôi sử dụng thuật ngữ “khách hàng doanh nghiệp” để chỉ đến doanh nghiệp – khách hàng của hệ thống chúng tôi và thuật ngữ “người dùng” để chỉ đến những người sử dụng website TMĐT của doanh nghiệp. Còn từ “khách hàng” nói chung sẽ hướng đến khách hàng của hệ thống chúng tôi đang đề cập đến.

- Tập dữ liệu sử dụng trong luận văn:
 - Dữ liệu mẫu tự sinh, tham khảo thông tin từ trang web vật dụng thú cưng Petmart, Cityzoo, Kikipet, bao gồm:
 - ✓ Cấu trúc thông tin về sản phẩm: mã định danh sản phẩm, tên sản phẩm, giá bán, hình ảnh và đường dẫn tới sản phẩm trên web TMĐT. Thông tin sản phẩm được sử dụng trong thuật toán content based.
 - ✓ Cấu trúc thông tin về khách hàng: mã định danh khách hàng, tài khoản đăng nhập, các thông tin cá nhân (số điện thoại, email, giới tính, tuổi, nơi ở). Thông tin cá nhân của khách hàng được sử dụng để xây dựng mô hình gợi ý sản phẩm sử dụng thuật toán demographic.
 - ✓ Cấu trúc hành vi khách hàng (thao tác click, scroll, mua hàng) trên các trang thương mại điện tử, lưu các thông tin về thời gian, định danh khách hàng, loại thao tác, sản phẩm được thao tác. Dựa trên dữ liệu hành vi khách hàng, chúng tôi xây dựng thuật toán collaborative filtering để gợi ý sản phẩm.
 - Sử dụng các thư viện trí tuệ nhân tạo có sẵn của Tensorflow, Sklearn, Gensim

1.4. Kết quả đạt được

Luận văn này đóng góp về mặt hệ thống gợi ý sản phẩm mà không tập trung vào việc cải thiện hiệu quả thuật toán. Quá trình thực hiện luận văn tạo điều kiện cho chúng tôi học thêm rất nhiều kỹ năng và kiến thức bổ ích từ kỹ năng mềm cho tới kiến thức về mặt lý thuyết cũng như thực hành, được cụ thể hóa như bên dưới:

- Về kỹ năng: Chúng tôi cải thiện được các kỹ năng như làm việc nhóm, lập kế hoạch, giải quyết vấn đề, quản lý thời gian, kỹ năng phân tích và truyền đạt ý tưởng. Những kỹ năng trên giúp chúng tôi phát triển bản thân và trở nên tự tin hơn trong học tập và công việc.
- Về lý thuyết: Chúng tôi hiểu được khái niệm SaaS, dịch vụ tương tác, các thuật toán gợi ý sản phẩm và tầm quan trọng của chúng hiện nay. Từ đó, chúng tôi ứng dụng vào thực tiễn thông qua việc xây dựng hệ thống gợi ý sản phẩm hướng dịch vụ. Điểm mới về lý thuyết, dựa trên Framework Trivi, chúng tôi định hướng hệ thống theo hướng dịch vụ có tương tác và tập trung vào lĩnh vực TMĐT.
- Về thực hành: Chúng tôi đã hoàn thành luận văn theo kế hoạch và mục tiêu đề ra. Chúng tôi đã khảo sát một số hệ thống hướng dịch vụ đang có trên thị trường để đánh giá nhằm rút kinh nghiệm cho hệ thống của chúng tôi. Đồng thời, chúng tôi xây dựng các website TMĐT phục vụ cho việc thực nghiệm và đã thành công tích hợp hệ thống gợi ý sản phẩm vào website TMĐT. Chúng tôi sẽ trình bày cụ thể hơn về quá trình triển khai hệ thống cũng như minh họa kết quả lần lượt ở các chương sau.

CHƯƠNG 2: NỀN TẢNG LÝ THUYẾT VÀ CÔNG NGHỆ

Chương 2 trình bày tổng quan về nền tảng lý thuyết và công nghệ được sử dụng trong luận văn. Trong chương này, đầu tiên chúng tôi trình bày định nghĩa và đánh giá vai trò của phần mềm hướng dịch vụ (SaaS – Software-as-a-Software) trong cuộc sống hiện nay. Từ đó, chúng tôi trình bày ý tưởng triển khai phần mềm gợi ý hướng dịch vụ và phân tích những ảnh hưởng mà nó mang lại. Sau cùng, chúng tôi trình bày các khái niệm về thuật toán gợi ý và giới thiệu công cụ, môi trường lập trình. Chương này được xem như là bước chuẩn bị kiến thức để thực hiện luận văn.

2.1. Phần mềm hướng dịch vụ (SaaS – Software-as-a-Service)

Phần này giới thiệu về phần mềm hướng dịch vụ (SaaS) và đánh giá của chúng tôi về nó. Từ đây, chúng tôi nhận ra tiềm năng phát triển SaaS và dựa trên SaaS, chúng tôi đề xuất triển khai phần mềm gợi ý hướng dịch vụ mà chúng tôi sẽ trình bày trong phần 2.2.

2.1.1. Tổng quan

SaaS (Software as a Service), là một mô hình kinh doanh phần mềm, trong đó một ứng dụng phần mềm được cung cấp dưới dạng dịch vụ cho người dùng thông qua internet. Thay vì phải cài đặt và quản lý phần mềm trên các thiết bị cá nhân hay máy chủ, người dùng có thể truy cập ứng dụng phần mềm thông qua trình duyệt web hoặc ứng dụng di động [12].

Hiện nay, SaaS ngày càng phổ biến trong lĩnh vực công nghệ. Theo báo cáo Thị trường phần mềm dịch vụ: Công nghệ và thị trường toàn cầu đến năm 2022, BCC Research dự đoán rằng ngành công nghiệp dựa trên mô hình SaaS được định giá \$44,4 tỷ vào năm 2017. Trong khi đó, người ta ước tính giá trị đó sẽ đạt được \$94,9 tỷ vào năm 2022, nghĩa là tốc độ tăng trưởng hàng năm

(CAGR) của thị trường SaaS vào khoảng 16,4%, chứng tỏ cơ hội mở rộng lớn cho ngành dịch vụ phần mềm này [13].

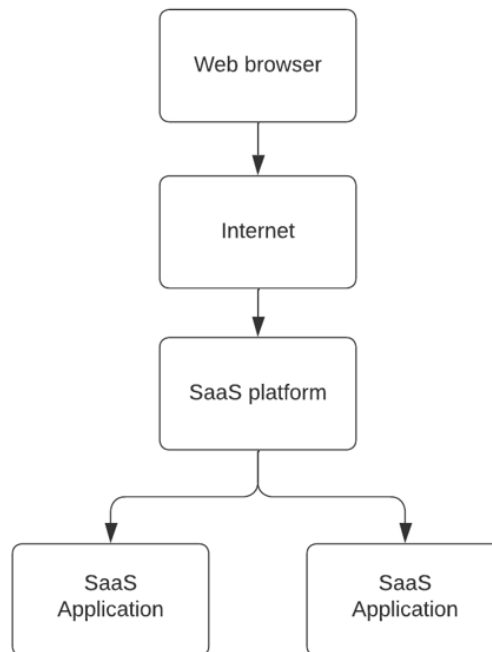
Một số phần mềm SaaS nổi tiếng hiện nay có thể kể đến Salesforce - hệ thống quản lý quan hệ khách hàng (CRM) hàng đầu trên thế giới, cung cấp các tính năng quản lý khách hàng, quản lý bán hàng, phân tích dữ liệu; Microsoft Office 365 - bộ công cụ văn phòng trực tuyến, cho phép người dùng truy cập các ứng dụng này từ bất kỳ thiết bị nào và lưu trữ dữ liệu trên đám mây; Zoom - ứng dụng họp trực tuyến cho phép người dùng trò chuyện, họp trực tuyến và hội nghị từ bất kỳ đâu.

SaaS (Software as a Service) và phần mềm truyền thống có một số sự khác biệt quan trọng sau đây:

| | SaaS | Phần mềm truyền thống |
|------------------------------|--|--|
| Model kinh doanh | Khách hàng doanh nghiệp có thể trả phí định kỳ (hàng tháng hoặc hàng năm). | Thường được bán dưới dạng sản phẩm với giá trọn đời hoặc giá trả trước một lần. |
| Cài đặt và triển khai | Thường được triển khai trên đám mây (cloud) và khách hàng doanh nghiệp có thể truy cập phần mềm thông qua trình duyệt. | Thường được cài đặt trên máy tính hoặc máy chủ của khách hàng doanh nghiệp. |
| Cập nhật phần mềm | Các cập nhật phần mềm thường được tự động thực hiện bởi nhà cung cấp dịch vụ. | Khách hàng doanh nghiệp thường phải tự cập nhật phần mềm và có thể mất nhiều thời gian và chi phí. |
| Quy mô | Thường phù hợp cho các khách hàng doanh nghiệp vừa và nhỏ, các cá nhân có nhu cầu sử dụng | Thường phù hợp cho các doanh nghiệp lớn với nhu cầu sử dụng |

| | | |
|-----------------------------------|--|--|
| | phần mềm trên một số máy tính hoặc thiết bị di động. | nhiều máy tính, máy chủ và mạng. |
| Bảo mật và quản lý dữ liệu | Nhà cung cấp dịch vụ thường chịu trách nhiệm quản lý và bảo mật dữ liệu của khách hàng doanh nghiệp. | Khách hàng doanh nghiệp phải tự quản lý và bảo mật dữ liệu của mình. |

Kiến trúc của SaaS được mô tả tổng quát qua Hình 2:



Hình 2. Kiến trúc SaaS

Người dùng sử dụng trình duyệt truy cập SaaS platform thông qua internet. SaaS platform chứa các ứng dụng SaaS khác nhau được cung cấp cho khách hàng doanh nghiệp. Mỗi ứng dụng SaaS có thể được quản lý và triển khai trên một máy

chủ riêng biệt. Các ứng dụng này được phát triển và quản lý bởi các nhà cung cấp SaaS và người dùng có thể sử dụng chúng thông qua trình duyệt.

Kiến trúc SaaS giúp đơn giản hóa việc quản lý và triển khai phần mềm, đồng thời cung cấp cho người dùng trải nghiệm sử dụng ứng dụng trực tuyến một cách tiện lợi và đơn giản.

2.1.2. Đánh giá chung

Khi sử dụng SaaS, việc quản lý ngân sách của doanh nghiệp trở nên dễ dàng hơn khi doanh nghiệp trả phí cho dịch vụ định kỳ và có thể chấm dứt dịch vụ bất cứ lúc nào. Ngoài ra, thay vì tải phần mềm và cài đặt, người dùng chỉ cần sử dụng thông qua trình duyệt trên Internet, do đó, mô hình SaaS giúp doanh nghiệp tiết kiệm được cả chi phí tiền mặt, thời gian, nhân lực, chi phí chuyển đổi và chi phí cơ hội.

SaaS cũng có những hạn chế nhất định, như việc phụ thuộc vào kết nối Internet và tính khả dụng của dịch vụ. Ngoài ra, việc lưu trữ dữ liệu trên đám mây có thể đặt ra nhiều vấn đề về bảo mật và quản lý dữ liệu. Tuy nhiên, SaaS vẫn là một lựa chọn tốt cho các doanh nghiệp và là một trong những hình thức cung cấp phần mềm dưới dạng dịch vụ nổi bật nhất hiện nay.

2.2. Phần mềm gợi ý hướng dịch vụ (RaaS – Recommender-as-a-Service)

Dựa trên kiến thức thu được về SaaS, chúng tôi nhận thấy rằng việc triển khai SaaS đem lại nhiều ảnh hưởng tích cực cho doanh nghiệp. Vì vậy, chúng tôi đề xuất phát triển một phần mềm gợi ý hướng dịch vụ. Phần trình bày này nhằm cung cấp nền tảng kiến thức cho luận văn của chúng tôi. Đây là trọng tâm của luận văn mà chúng tôi nghiên cứu.

2.2.1. Dịch vụ tương tác (Interactive service)

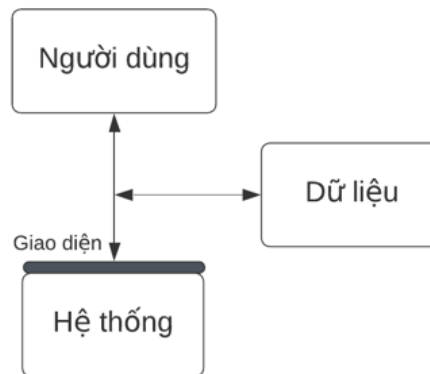
❖ Tổng quan

Dịch vụ tương tác là một dạng dịch vụ tương tác giữa người dùng và hệ thống, trong đó người dùng có thể tương tác với hệ thống thông qua các giao diện tương tác, như website, ứng dụng di động hoặc thiết bị IoT, điều này cho phép người dùng có thể tương tác trực tiếp với hệ thống, thực hiện các tác vụ và nhận phản hồi từ hệ thống ngay lập tức.

Dịch vụ tương tác tập trung vào sự tương tác của người dùng với doanh nghiệp, ở đây chúng tôi muốn nói đến sự tương tác của người dùng với trang web thương mại điện tử (TMĐT). Sự tương tác của người dùng ảnh hưởng đến chất lượng dịch vụ và kết quả hành vi cũng như hiệu quả kinh doanh.

❖ Mô hình dịch vụ tương tác

Hình 3 minh họa tổng quan mô hình dịch vụ tương tác, trong đó, mô hình có ba thành phần chính:



Hình 3. Mô hình dịch vụ tương tác

- Người dùng: Là những người sử dụng dịch vụ và tương tác với hệ thống thông qua giao diện. Người dùng có thể yêu cầu thông tin hoặc thực hiện một số hành động cụ thể.
- Hệ thống: Là một phần mềm hoặc nền tảng được sử dụng để cung cấp dịch vụ. Hệ thống bao gồm các thành phần phần mềm và cơ sở hạ tầng để quản lý dữ liệu, xử lý yêu cầu của người dùng và cung cấp phản hồi.

- Dữ liệu: Là các thông tin được thu thập từ người dùng hoặc từ các nguồn khác để cung cấp cho người dùng thông tin và giải pháp tốt nhất. Dữ liệu bao gồm thông tin về người dùng, sản phẩm, hoạt động của người dùng, đánh giá, phản hồi và các thông tin liên quan khác.

Với mô hình này, người dùng tương tác với hệ thống bằng cách gửi yêu cầu hoặc cung cấp thông tin, sau đó hệ thống sẽ xử lý yêu cầu hoặc cung cấp phản hồi dựa trên thông tin có sẵn trong hệ thống và dữ liệu được thu thập.

❖ Hình thức dịch vụ tương tác

Với sự phát triển của công nghệ, các hình thức của dịch vụ tương tác càng ngày càng phổ biến và đa dạng hơn. Nó đã trở thành một phần không thể thiếu trong việc xây dựng mối quan hệ với người dùng, mang lại lợi ích lớn cho cả người dùng và doanh nghiệp. Có thể kể đến hai hình thức dịch vụ tương tác chính sau:

- Người dùng chủ động yêu cầu hỗ trợ từ hệ thống: Đây là hình thức truyền thống, khi có nhu cầu hoặc vấn đề về sản phẩm, ví dụ như người dùng sử dụng công cụ tìm kiếm mà hệ thống cung cấp hoặc chat để nhận hỗ trợ.
- Hệ thống tự động phục vụ người dùng: Hệ thống nhận biết nhu cầu và chủ động đáp ứng bằng cách phân tích thao tác của người dùng trên hệ thống.

Có thể hiểu như sau, với hình thức truyền thống, khi muốn mua các sản phẩm nào đó, người dùng có thể tìm kiếm theo tên, loại sản phẩm trên thanh tìm kiếm, hệ thống sẽ phản hồi theo yêu cầu cụ thể được đưa ra. Tuy nhiên, với luận văn này, chúng tôi muốn cung cấp hệ thống có thể phân tích thao tác của người dùng (click, scroll, v.v) để chủ động phản hồi nhằm đáp ứng nhu cầu. Với hình thức này, người dùng sẽ cảm thấy được sự thấu hiểu, điều đó làm tăng độ hài lòng đối với hệ thống.

❖ Đặc trưng dịch vụ tương tác

Dịch vụ tương tác đòi hỏi hệ thống phải có khả năng xử lý tương tác nhanh và hiệu quả, đảm bảo rằng người dùng có thể tương tác với hệ thống một cách dễ dàng và không gặp phải những gián đoạn hoặc trễ truyền. Tuy nhiên, nó cũng đem lại nhiều lợi ích cho người dùng và doanh nghiệp, bao gồm tăng tính tương tác và thân thiện với người dùng, giảm chi phí tương tác và tăng tính năng suất.

Các đặc trưng chính của dịch vụ tương tác bao gồm:

- Tương tác trực tiếp với người dùng: Dịch vụ tương tác cho phép người dùng tương tác với hệ thống thông qua giao diện, giúp cải thiện trải nghiệm người dùng.
- Đa dạng giao diện người dùng: Dịch vụ tương tác hỗ trợ nhiều loại giao diện người dùng, bao gồm ứng dụng di động, web và các giao diện khác, đáp ứng nhu cầu sử dụng của người dùng ở nhiều nền tảng khác nhau.
- Tính linh hoạt: Dịch vụ tương tác có thể thích ứng với các yêu cầu khác nhau của người dùng tùy thuộc vào nhu cầu của người dùng.
- Tự động hóa: Dịch vụ tương tác có thể sử dụng các công nghệ tự động hóa (học máy, tổ chức quy trình tự động) để cung cấp các dịch vụ và giải quyết các yêu cầu của người dùng.
- Độ tin cậy cao: Dịch vụ tương tác đảm bảo phản hồi nhanh chóng với yêu cầu của người dùng.
- **Công nghệ sử dụng**

Tất cả các đặc trưng trên đóng vai trò quan trọng trong việc cung cấp trải nghiệm tốt cho người dùng và đáp ứng nhu cầu của họ. Tuy nhiên, để cung cấp dịch vụ tương tác hiệu quả, cần có sự trợ giúp từ công nghệ thông tin, các công nghệ này bao gồm:

- Phần mềm tương tác: Các phần mềm này tạo ra môi trường tương tác, cung cấp tính năng phản hồi nhằm đáp ứng nhu cầu người dùng. Trong luận văn

này, phần mềm tương tác nói đến các website TMDT, hỗ trợ việc ghi nhận tương tác của người dùng trên website.

- Hệ thống quản lý tương tác người dùng: Hệ thống lưu trữ thông tin và theo dõi các tương tác của người dùng, phân tích dữ liệu người dùng. Đây chính là hệ thống mà luận văn hướng đến, chúng tôi áp dụng thuật toán để phân tích tương tác người dùng nhằm đưa ra gợi ý sản phẩm, giúp doanh nghiệp TMDT tạo ra các chiến lược tương tác và tăng cường trải nghiệm người dùng.
- Máy học (Machine Learning): AI (Trí tuệ nhân tạo) là yếu tố cốt lõi để tạo nên dịch vụ hiệu quả. Để phân tích dữ liệu người dùng, cần sử dụng thuật toán máy học phù hợp, trong đó, đối với việc xây dựng dịch vụ gợi ý sản phẩm, chúng tôi sử dụng các thuật toán sẽ được trình bày chi tiết ở phần 2.3.
- Ngoài ra, xử lý ngôn ngữ tự nhiên (NLP) được sử dụng để phân tích các dữ liệu người dùng từ các phản hồi, đánh giá sản phẩm, dịch vụ. Luận văn này chúng tôi không đề cập đến việc gợi ý sản phẩm dựa trên lời văn đánh giá của người dùng mà chỉ tập trung vào các sự kiện tương tác (click, scroll, thêm giỏ hàng, bỏ khỏi giỏ hàng).

- **Đánh giá chung**

Trên hệ thống bán hàng điện tử, sự tương tác của người dùng rất quan trọng vì nó có ảnh hưởng lớn trong việc xây dựng mối quan hệ giữa người dùng và doanh nghiệp. Tương tác của người dùng cũng cung cấp cho doanh nghiệp thông tin quan trọng về nhu cầu và mong muốn của người dùng, giúp doanh nghiệp có thể cải thiện sản phẩm và dịch vụ của mình để đáp ứng các nhu cầu của họ. Có thể nhận thấy rằng, từ sự tương tác của người dùng, doanh nghiệp có thể phân tích hành vi và đưa ra phương án kinh doanh phù hợp. Dịch vụ tương tác dần trở thành mối quan tâm lớn trên thị trường. Hiện nay, công nghệ đang phát triển

nhanh chóng, những xu hướng mới vẫn đang được hình thành, doanh nghiệp cần phản ứng kịp thời với những thay đổi về xu hướng, với thói quen mua sắm, với hành vi của người dùng. Việc đầu tư vào các công cụ tương tác người dùng là một đầu tư quan trọng đối với các doanh nghiệp hoạt động trong lĩnh vực bán hàng điện tử.

Từ những tìm hiểu trên, nhận ra khả năng phát triển của mô hình tương tác hướng dịch vụ, chúng tôi nghiên cứu và triển khai mô hình tư vấn hướng dịch vụ để đưa ra cái nhìn tổng quan kèm theo thực nghiệm minh họa nhằm đánh giá cho sự phát triển của mô hình trong tương lai. Đây cũng là bài toán mà chúng tôi sẽ giải quyết trong luận văn này.

2.2.2. Phần mềm gợi ý hướng dịch vụ

Đầu tiên, chúng tôi xin trình bày về phần mềm gợi ý sản phẩm. Phần mềm gợi ý là loại phần mềm có khả năng phân tích, hiểu được sở thích, nhu cầu và hành vi của người dùng để đưa ra các gợi ý sản phẩm hay dịch vụ phù hợp. Phần mềm này thường được sử dụng trong các lĩnh vực như thương mại điện tử, du lịch, giải trí, dịch vụ và nhiều lĩnh vực khác. Với sự phát triển của công nghệ và sự xuất hiện của nhiều hệ thống dịch vụ trực tuyến, phần mềm gợi ý đóng vai trò quan trọng trong việc giúp người dùng tìm kiếm các dịch vụ phù hợp với nhu cầu của họ.

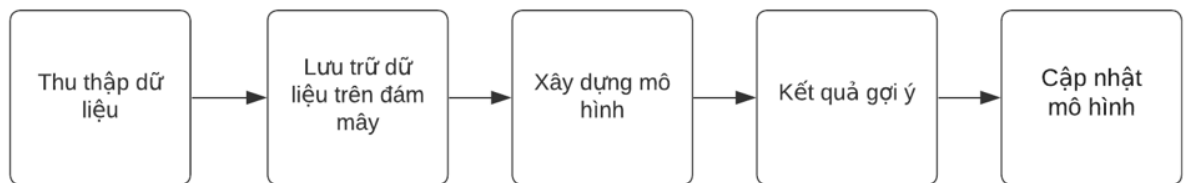
Một số ứng dụng nổi tiếng có cài đặt phần mềm gợi ý như Netflix – đề xuất phim và chương trình truyền hình phù hợp với sở thích khách hàng; Amazon – đề xuất sản phẩm phù hợp với lịch sử mua hàng và đánh giá của khách hàng; YouTube - đề xuất cho người dùng các video và kênh phù hợp với sở thích và lịch sử xem của họ.

Có thể hiểu về phần mềm gợi ý hoạt động theo cách như sau, khi một người dùng truy cập trang web bán hàng, phần mềm gợi ý hướng dịch vụ sẽ phân tích lịch sử mua hàng, các sản phẩm đã xem và các sản phẩm liên quan để đưa ra các

gợi ý sản phẩm phù hợp với nhu cầu của người dùng. Tương tự, trong lĩnh vực du lịch, phần mềm này có thể đưa ra các gợi ý về các điểm đến, khách sạn, hoạt động giải trí và ẩm thực phù hợp với sở thích và nhu cầu của khách hàng.

Nói về đề xuất phần mềm gợi ý hướng dịch vụ, chúng tôi triển khai dựa trên ý tưởng xây dựng phần mềm gợi ý, tuy nhiên, các chức năng chính sẽ được chúng tôi đóng gói và cung cấp cho các doanh nghiệp vừa và nhỏ dưới dạng phần mềm dịch vụ. Chúng tôi sẽ trình bày cụ thể hơn các chức năng chính này trong chương 3.

Quy trình của phần mềm gợi ý hướng dịch vụ bao gồm các công việc chính như hình 4:



Hình 4. Quy trình gợi ý hướng dịch vụ

- Thu thập dữ liệu: Phần mềm sẽ thu thập dữ liệu về thông tin người dùng, bao gồm lịch sử sử dụng dịch vụ, mục đích sử dụng, địa điểm, thời gian, số lượng người sử dụng, đánh giá, v.v. Dữ liệu này sẽ được lưu trữ trong cơ sở dữ liệu.
- Công nghệ điện toán đám mây: Do hệ thống gợi ý sản phẩm phát triển theo hướng dịch vụ, chúng tôi sử dụng điện toán đám mây để lưu trữ dữ liệu trên đám mây, cho phép người dùng được truy cập hệ thống từ bất kỳ thiết bị có kết nối Internet.
- Xây dựng mô hình: Dựa trên dữ liệu đã thu thập, phần mềm sẽ sử dụng các kỹ thuật máy học (machine learning) để xây dựng mô hình gợi ý sản phẩm.

Mô hình này sẽ phân tích dữ liệu và đưa ra các đề xuất dịch vụ phù hợp cho người dùng.

- Gợi ý dịch vụ: Dựa trên mô hình gợi ý sản phẩm, phần mềm sẽ gợi ý các sản phẩm, dịch vụ phù hợp với nhu cầu của người dùng. Các đề xuất này có thể được hiển thị trên giao diện của phần mềm hoặc tùy theo phương án kinh doanh của doanh nghiệp (gửi đến người dùng qua email hoặc tin nhắn).
- Cập nhật mô hình: Phần mềm sẽ liên tục cập nhật mô hình gợi ý dựa trên dữ liệu mới để đưa ra các gợi ý dịch vụ ngày càng chính xác và phù hợp với nhu cầu của người dùng.

Phần mềm gợi ý theo hướng dịch vụ giúp tăng tính trải nghiệm của người dùng, giảm thời gian tìm kiếm và tăng tính tiện lợi trong việc tìm kiếm và chọn lựa sản phẩm, dịch vụ. Đối với doanh nghiệp, phần mềm này còn giúp họ tăng doanh số bằng cách đưa ra các gợi ý sản phẩm và dịch vụ phù hợp với nhu cầu của người dùng và giảm thiểu chi phí đáng kể nhờ vào việc triển khai phần mềm theo hướng dịch vụ.

2.3. Nghiên cứu thuật toán gợi ý

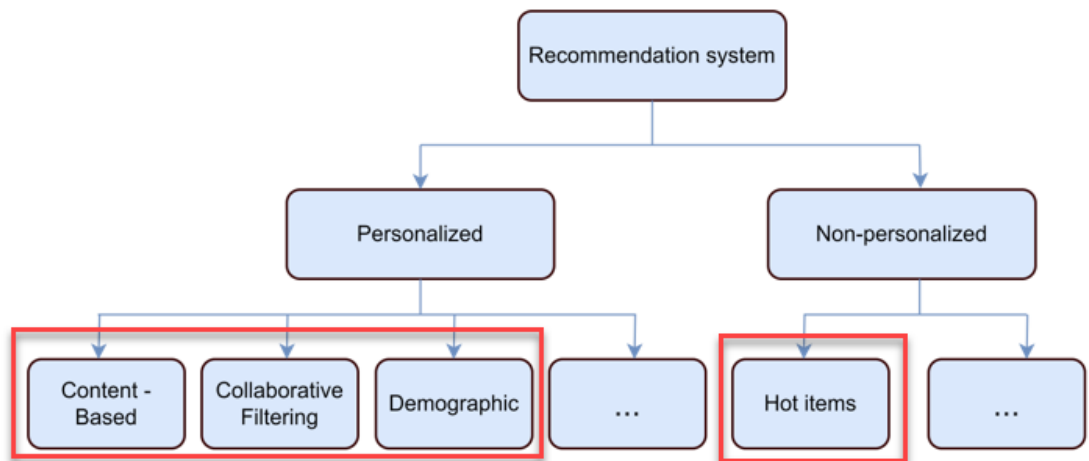
Sau khi chuẩn bị các kiến thức liên quan đến mô hình phát triển, chúng tôi sẽ trình bày về các thuật toán gợi ý sản phẩm được áp dụng trong luận văn. Để đưa ra hướng tiếp cận thuật toán phù hợp, chúng tôi đã trải nghiệm và đánh giá tình trạng chung tại các website TMĐT và liệt kê các thư viện cung cấp các thuật toán mà chúng tôi sử dụng.

Thuật toán gợi ý sản phẩm (Product Recommendation) là một trong những phương pháp quan trọng trong lĩnh vực khai thác dữ liệu, giúp đề xuất các sản phẩm tương tự hoặc liên quan đến sản phẩm mà người dùng có thể quan tâm. Thuật toán gợi ý tốt có thể giúp doanh nghiệp tăng doanh số, nâng cao trải nghiệm

mua sắm của người dùng. Chúng tôi minh họa phân loại các thuật toán gợi ý sản phẩm qua hình 5.

Các thuật toán gợi ý sản phẩm sử dụng dữ liệu như lịch sử mua hàng, thông tin cá nhân, đánh giá của người dùng, lượt xem sản phẩm, sản phẩm liên quan để tạo ra những gợi ý sản phẩm phù hợp với nhu cầu của người dùng. Thuật toán gợi ý có 2 nhánh chính là cá nhân hóa và không cá nhân hóa.

- Thuật toán gợi ý hướng cá nhân hóa đánh mạnh vào việc gợi ý cho từng đối tượng người dùng. Các thuật toán điển hình thường dùng như là Gợi ý theo nội dung (Content-based), Lọc cộng tác (Collaborative filtering), Nhân khẩu học (Demographic), v.v. Đặc điểm chung của các thuật toán hướng cá nhân hóa là cần có dữ liệu về người dùng.
- Thuật toán gợi ý hướng không cá nhân hóa đánh mạnh vào việc gợi ý cho đại chúng. Những thuật toán gợi ý truyền thông có thể kể đến như gợi ý sản phẩm bán chạy nhất, gợi ý sản phẩm mới nhất, v.v.



Hình 5. Phân loại thuật toán gợi ý sản phẩm [14]

Trong luận văn này, chúng tôi tập trung xây dựng dịch vụ hệ thống gợi ý cho các trang TMĐT vừa và nhỏ. Đối với các trang TMĐT này, tính đa dạng về thuật toán gợi ý là không thể thiếu. Do đó, trong phần dưới đây chúng tôi lần lượt

mô tả và phân tích các phương thức gợi ý được cài đặt dưới dạng các dịch vụ là Gợi ý theo nội dung (Content-based), Lọc cộng tác (Collaborative filtering), Nhân khẩu học (Demographic), Gợi ý sản phẩm phổ biến (Hot items).

2.3.1. Lọc cộng tác (Collaborative Filtering)

❖ Giới thiệu

Collaborative filtering là một thuật toán gợi ý sản phẩm phổ biến hiện nay. Thuật toán này sử dụng chủ yếu là các đánh giá sản phẩm của tập người dùng để tìm kiếm các sản phẩm mà người dùng mục tiêu đã đánh giá cao.

Thuật toán collaborative filtering có hai loại chính: Dựa trên người dùng (user-based) và dựa trên sản phẩm (item-based). Trong user-based collaborative filtering, hệ thống đưa ra gợi ý dựa trên sự tương đồng giữa người dùng, trong khi item-based collaborative filtering sẽ đưa ra gợi ý dựa trên sự tương đồng giữa các sản phẩm. Thuật toán collaborative filtering được sử dụng rộng rãi trong các hệ thống gợi ý sản phẩm, nhất là trong các trang TMĐT [15].

❖ Tình trạng tại các trang TMĐT vừa và nhỏ

Sau khi thử trải nghiệm một số trang thương mại điện tử vừa và nhỏ, chúng tôi nhận thấy rằng người dùng hiếm khi để lại bình luận nhận xét sau khi mua hàng. Từ đó, việc dựa trên đánh giá của người dùng để đưa ra gợi ý sẽ không mang lại hiệu quả cao do thiếu hụt dữ liệu, vì vậy chúng tôi đưa ra hướng tiếp cận đến hành vi người dùng trên trang TMĐT.

❖ Hướng tiếp cận

Thông qua khảo sát ở trên, các trang TMĐT thiếu hụt dữ liệu đánh giá của người dùng. Để giải quyết vấn đề đó, chúng tôi đề xuất góc nhìn thay thế các dữ liệu đánh giá của người dùng bằng hành vi của người dùng, ví dụ như lướt xem sản phẩm, bấm vào sản phẩm, thêm sản phẩm vào giỏ hàng, xóa sản phẩm khỏi giỏ hàng, v.v. Từ các dữ liệu đó, chúng tôi nội suy ra đánh giá người dùng.

❖ Yêu cầu dữ liệu

Với hướng tiếp cận đó, các loại đối tượng dữ liệu cần thu thập cần có bối cảnh như sau:

Khi người dùng sử dụng trang TMĐT, chúng tôi chỉ quan tâm đến tương tác của họ thực hiện trên sản phẩm nào (theo định danh – id) mà không cần xét đến thông tin của sản phẩm cũng như thông tin cá nhân của người dùng.

❖ Thư viện sử dụng

Hiện nay, Tensorflow đã có cung cấp bộ gợi ý sản phẩm trong thư viện của mình. Dựa trên thư viện của Tensorflow, chúng tôi sử dụng theo mô hình neural network 2 towers đã được cung cấp sẵn.

2.3.2. Dựa trên nhân khẩu học (Demographic Filtering)

❖ Giới thiệu

Demographic Filtering là một trong những thuật toán gợi ý sản phẩm dựa trên đặc tính thông tin cơ bản của người dùng như độ tuổi, giới tính, nghề nghiệp, v.v và sử dụng chúng để tìm các sản phẩm phù hợp với nhu cầu của người dùng đó. Về cơ bản, thuật toán này giả định rằng những người cùng nhóm nhân khẩu học có xu hướng có sở thích giống nhau, do đó sản phẩm hoặc dịch vụ được gợi ý sẽ giống nhau.

Thuật toán demographic là một phương pháp đơn giản và dễ thực hiện, tuy nhiên nó cũng có một số hạn chế như khả năng gợi ý sản phẩm bị hạn chế trong cùng một nhóm nhân khẩu học, sự khác biệt giữa các người dùng cùng nhóm nhân khẩu học vẫn có thể khá lớn.

❖ Tình trạng tại các trang TMĐT vừa và nhỏ

Các trang TMĐT vừa và nhỏ đều có phần thu thập những thông tin cá nhân cơ bản như độ tuổi, giới tính, thông qua phần đăng ký tài khoản; thu thập dữ liệu vị trí thông qua địa chỉ giao hàng của đơn đặt hàng.

❖ Hướng tiếp cận

Với tính chất dữ liệu cá nhân người dùng trên, chúng tôi sử dụng thuật toán gom cụm Kmeans để phân cụm người dùng. Sau đó sẽ gợi ý sản phẩm nổi bật nhất trong cụm cho người dùng thuộc cụm tương ứng

❖ **Yêu cầu dữ liệu**

Dữ liệu cá nhân của người dùng, điển hình như là (tuổi, vị trí địa lý, giới tính).

❖ **Thư viện sử dụng**

Sklearn là một trong những thư viện cung cấp thuật toán gom cụm mạnh mẽ, trong đó có Kmeans.

2.3.3. Dựa trên nội dung (Content-based)

❖ **Giới thiệu**

Content-based là thuật toán gợi ý sản phẩm phù hợp với mục đích của người dùng. Nó hoạt động dựa trên thông tin về nội dung của sản phẩm, như tên, mô tả, thể loại, giá, v.v. để tìm ra được những sản phẩm tương đồng với sản phẩm đang được xem bởi người dùng. Điều này có nghĩa là, thuật toán sẽ dựa trên các thuộc tính và đặc trưng của sản phẩm để đưa ra gợi ý cho người dùng [16].

❖ **Tình trạng tại các trang TMĐT vừa và nhỏ**

Content-based là một thuật toán khả quan khi các trang thương mại điện tử đã có đầy đủ các thông tin của sản phẩm.

❖ **Hướng tiếp cận**

Thuật toán content-based không yêu cầu dữ liệu đánh giá người dùng và có khả năng gợi ý các sản phẩm tương tự với sản phẩm mà người dùng vừa tương tác trên hệ thống. Thuật toán này có thể áp dụng được cho người dùng mới của web TMĐT khi chưa có bất cứ dữ liệu nào về người dùng mới.

❖ **Yêu cầu dữ liệu**

Thông tin chi tiết của sản phẩm như (Tên sản phẩm, Giá sản phẩm, Loại sản phẩm) và thông tin về tương tác của người dùng tại thời điểm hiện tại trên trang TMĐT.

❖ **Thư viện sử dụng**

Gensim là một thư viện Python cho phép tìm kiếm và xử lý từ vựng và văn bản trên lượng lớn dữ liệu với các thuật toán xử lý từ vựng và văn bản. Gensim được sử dụng rộng rãi trong các mô hình xử lý từ vựng và văn bản, đặc biệt là các mô hình gợi ý nội dung.

2.3.4. Gợi ý sản phẩm phổ biến nhất (Hot items)

❖ **Giới thiệu**

Bên cạnh các thuật toán về máy học mạnh mẽ, các thuật toán như gợi ý sản phẩm nổi bật nhất, sản phẩm mới nhất hiện nay vẫn được áp dụng tốt trong lĩnh vực bán hàng mặc dù thuật toán này hoạt động khá đơn giản.

❖ **Hướng tiếp cận**

Có rất nhiều định nghĩa về độ phổ biến của sản phẩm, ví dụ như được mua nhiều nhất, doanh thu nhiều nhất, v.v. Ở luận văn này, chúng tôi định nghĩa sản phẩm phổ biến nhất là sản phẩm có lượt tương tác nhiều nhất.

❖ **Yêu cầu dữ liệu**

Thông tin lượt tương tác giữa người dùng và sản phẩm.

2.3.5. Kết luận

Sau khi tìm hiểu và nghiên cứu các thuật toán gợi ý sản phẩm, chúng tôi rút ra được kết luận như sau:

Đối với các thuật toán trên, cần quan tâm đến các kỹ thuật tối ưu hóa để tăng tốc độ tính toán. Các thuật toán nhận đầu vào là dữ liệu từ cơ sở dữ liệu và đầu ra là một danh sách các sản phẩm, sau đó hệ thống sẽ chuyển kết quả này đến người dùng. Chúng tôi sẽ đề cập cách khách hàng doanh nghiệp xử lý kết quả này trong phần 3.3.1.

Thuật toán gợi ý sản phẩm mang lại nhiều lợi ích cho người dùng và doanh nghiệp. Nhìn chung, thuật toán gợi ý sản phẩm hiệu quả, phù hợp sẽ giúp doanh nghiệp có thể tối ưu doanh số bán hàng và cải thiện trải nghiệm người dùng. Tuy

nhien, chúng tôi chỉ triển khai thuật toán ở mức thuật toán thích nghi được với hệ thống và không bàn luận về độ hiệu quả trong luận văn này.

2.4. Công cụ và môi trường phát triển

Với mục đích có thể xây dựng được hệ thống dịch vụ tương tác, ở phần này chúng tôi giới thiệu những công cụ và môi trường phát triển của luận văn. Các công cụ và môi trường trình bày sau đây dựa trên kiến trúc hệ thống hệ thống mà chúng tôi sẽ đề cập trong phần 3.3.

Với tốc độ phát triển nhanh chóng của công nghệ, những công cụ và môi trường phát triển được thay đổi chóng mặt. Trong luận văn này, chúng tôi sử dụng một số công cụ đã và đang có chỗ đứng trên thị trường về độ phổ biến với mục đích xây dựng và triển khai một hệ thống dịch vụ gợi ý có tương tác. Các thành phần chính bao gồm: giao diện cho admin của các trang TMĐT (Front-end), hệ thống phụ trợ cung cấp các API (Back-end), bộ lập lịch (Scheduler), cơ sở dữ liệu (Database).

2.4.1. Hệ thống phụ trợ – Django

Django là một Framework lập trình Web bậc cao, mã nguồn mở được viết bằng ngôn ngữ lập trình Python. Django được lập trình với mô hình xác thực người dùng tối ưu, phù hợp với các dự án cần có độ bảo mật cao. Đặc biệt, Django có nguồn tài liệu hướng dẫn tốt, tuân theo nguyên tắc DRY và hỗ trợ ORM (Object Relational Mapping).

Trong luận văn này, chúng tôi sử dụng Django là một hệ thống phụ trợ, để tạo API, quản lý dữ liệu dễ dàng và nhanh chóng nhờ vào những modules đã được cài đặt sẵn: biểu mẫu, xác thực người dùng, quản trị web. Với nền ngôn ngữ là Python, nên Django rất dễ dàng trong việc tích hợp, sử dụng các thư viện cung cấp thuật toán chúng tôi đã nói ở phần 2.3.

2.4.2. Giao diện khách hàng doanh nghiệp – ReactJS

ReactJS là một thư viện Javascript dùng để xây dựng giao diện người dùng, sử dụng mô hình dữ liệu trực quan (Virtual DOM) để tối ưu hóa việc cập nhật giao diện. ReactJS hỗ trợ tạo những thành phần UI có tính tương tác cao và có thể tái sử dụng.

Trong luận văn này, chúng tôi sử dụng ReactJS để tạo giao diện trang quản lý cho quản trị viên của các website thương mại điện tử. Trang quản lý này sẽ bao gồm các thông tin liên quan đến thống kê các sự kiện diễn ra trên website chính của họ, đồng thời cho phép quản trị viên cài đặt các chiến lược gợi ý phù hợp theo nhu cầu.

2.4.3. Các cơ sở dữ liệu quan hệ

Cơ sở dữ liệu quan hệ (Relational Database Management System – RDBMS) dựa trên mô hình dữ liệu quan hệ, sử dụng các bảng để lưu trữ dữ liệu và quan hệ giữa các bảng, trong đó, các bảng được liên kết bằng các khóa chính và khóa ngoại, giúp tránh trùng lặp dữ liệu và đảm bảo tính toàn vẹn của dữ liệu. RDBMS được sử dụng rộng rãi nhờ các tính năng bảo mật, dễ quản lý, cung cấp công cụ và ngôn ngữ truy vấn dữ liệu để lọc, tìm kiếm và sửa đổi dữ liệu một cách dễ dàng.

Dịch vụ gợi ý mà chúng tôi nghiên cứu trong khóa luận này có thể đáp ứng kết nối được hầu hết cơ sở dữ liệu quan hệ nổi tiếng ngày nay, điển hình như PostgreSQL, MySQL, Oracle, MicrosoftSQL.

Trong luận văn này, các cơ sở dữ liệu được dùng để lưu 2 loại dữ liệu:

- Dữ liệu do quản trị viên của trang TMĐT cung cấp.
- Dữ liệu được bắt từ người dùng của trang TMĐT

2.4.4. Bộ lập lịch – Airflow

Airflow là một nền tảng quản lý workflows phổ biến được sử dụng để quản lý và thực hiện các tác vụ liên quan đến Big Data và Machine Learning. Nó cho

phép thiết lập và quản lý các workflows được chạy theo lịch trình, tự động hoặc thủ công. Airflow còn cung cấp nhiều tính năng hỗ trợ thống kê và theo dõi, gửi email, giúp đánh giá hiệu suất của tác vụ đang chạy.

Trong luận văn này, chúng tôi sử dụng Airflow để hỗ trợ cho việc lập lịch, cài đặt thời gian huấn luyện mô hình gợi ý sản phẩm theo yêu cầu bên phía doanh nghiệp.

2.4.5. Điện toán đám mây – Digital Ocean

Digital Ocean là một nhà cung cấp dịch vụ đám mây (cloud computing), cung cấp các dịch vụ máy chủ ảo (VPS), cho phép người dùng tạo, quản lý và triển khai các ứng dụng trên đám mây của họ. Digital Ocean được đánh giá cao nhờ vào độ tin cậy, tốc độ và dễ sử dụng. Ngoài ra, Digital Ocean cũng cung cấp các tính năng như bảo mật, sao lưu dữ liệu và giám sát hệ thống.

Trong luận văn này, chúng tôi lưu trữ dữ liệu và triển khai hệ thống gợi ý sản phẩm trên đám mây nhờ vào Digital Ocean.

2.5. Kết luận chương

Chúng tôi đã trình bày tổng quan về nền tảng lý thuyết và công nghệ phục vụ cho luận văn. Với những kiến thức trên, chúng tôi sẽ bắt đầu thực hiện các bước như khảo sát, phân tích yêu cầu đến lập trình và triển khai những nền tảng lý thuyết này để xây dựng hệ thống dịch vụ gợi ý có tương tác lẫn lượt ở các chương sau.

CHƯƠNG 3: MÔ HÌNH TRIỂN KHAI DỊCH VỤ GỢI Ý

Chương 3 trình bày chi tiết về mô hình gợi ý sản phẩm hướng dịch vụ mà chúng tôi đã đề xuất trong luận văn. Từ kiến thức đã nghiên cứu ở chương 2, trong chương 3, trước hết, chúng tôi sẽ khảo sát hiện trạng các dịch vụ gợi ý hiện tại, từ đó phân tích để xác định các yêu cầu của một hệ thống gợi ý sản phẩm hướng dịch vụ. Giải pháp của chúng tôi được triển khai dựa trên Framework Trivi, đã được công bố trong bài báo Trivi: A conceptual Framework for Customer Intelligence Systems for Small and Medium-sized Enterprises [6]. Nguyên tắc cơ bản và mô hình dữ liệu được đề xuất từ hệ thống Trivi cũng sẽ được giới thiệu sơ lược trong chương này. Sau cùng, chúng tôi sẽ trình bày giải pháp chúng tôi đề xuất cho mô hình triển khai dịch vụ gợi ý dựa trên nền tảng Framework Trivi, bao gồm kiến trúc hệ thống, mô hình cơ sở dữ liệu và sơ đồ tương tác (sequence diagram) của các thành phần trong hệ thống.

3.1. Phân tích yêu cầu

Phần này nhằm khảo sát các hệ thống dịch vụ tương tác và các hệ thống gợi ý sản phẩm hiện nay để phân tích và đưa ra đánh giá các hệ thống này. Từ đó, chúng tôi đưa ra hướng triển khai hệ thống phù hợp cho các doanh nghiệp vừa và nhỏ.

3.1.1. Khảo sát hệ thống thực tế

Hệ thống dịch vụ tương tác hiện nay ngày càng phát triển đa dạng để đáp ứng nhu cầu của người dùng. Chúng tôi đã thực hiện khảo sát trên một số hệ thống dịch vụ tương tác cũng như hệ thống gợi ý sản phẩm và rút ra được các kết luận như sau:

a. Hệ thống dịch vụ tương tác:

❖ Hệ thống quản lý khách hàng (CRM)

CRM (Customer Relationship Management) là một hệ thống quản lý khách hàng, được sử dụng để quản lý và tối ưu quá trình tương tác của doanh nghiệp với khách hàng. Theo một khảo sát của Super Office, 92% doanh nghiệp nói rằng CRM là công cụ quan trọng trong việc đáp ứng mục tiêu bán hàng [17]. Theo phần 2.2.1 trình bày về hình thức dịch vụ tương tác, CRM đang cung cấp dịch vụ cho khách hàng theo hình thức 1, nghĩa là CRM chỉ có nhiệm vụ cung cấp công cụ để khách hàng tương tác với giao diện nhằm thực hiện các chức năng có sẵn. Trên giao diện làm việc của CRM, doanh nghiệp có thể lưu trữ thông tin khách hàng, tạo lịch trình và theo dõi các hoạt động với khách hàng như hẹn gặp hoặc gửi thư cảm ơn. Dựa vào khả năng phân tích của CRM, doanh nghiệp có thể tối ưu hóa chiến lược kinh doanh, phát triển sản phẩm và nâng cao chất lượng dịch vụ.

Chi phí mua một hệ thống CRM có thể dao động tùy thuộc vào nhiều yếu tố khác nhau như kích thước của doanh nghiệp, tính năng của hệ thống, và các yêu cầu kỹ thuật khác. Tổng chi phí để mua một hệ thống CRM phụ thuộc vào nhiều yếu tố và có thể dao động từ vài nghìn đến vài trăm nghìn USD cho các gói dịch vụ khác nhau. Việc triển khai và sử dụng hệ thống CRM cần đòi hỏi một số kỹ năng và kiến thức chuyên môn, đồng thời cần phải đầu tư vào hệ thống phần mềm, đào tạo nhân viên và quản lý các thông tin khách hàng một cách chặt chẽ. Hệ thống CRM có thể không tương thích hoặc khó tích hợp với các hệ thống khác mà doanh nghiệp đang sử dụng. Nếu không sử dụng đúng cách, hệ thống CRM có thể gây ra những rắc rối và chi phí không cần thiết cho doanh nghiệp.

❖ Chatbot

Chatbot là một công nghệ được phát triển để giúp tương tác với người dùng qua trò chuyện trực tuyến. Chưa có khảo sát về tỉ lệ doanh nghiệp sử dụng chatbot vào năm 2023, tuy nhiên theo những xu hướng hiện tại, chatbot đang được coi là một công nghệ tiềm năng để cải thiện trải nghiệm khách hàng và nâng cao hiệu

quả kinh doanh của các doanh nghiệp. Theo Gartner, vào năm 2022, khoảng 70% các doanh nghiệp sẽ sử dụng ít nhất một chatbot hoặc ứng dụng tương tự để tương tác với khách hàng của họ [18]. Ngoài ra, Juniper Research dự báo rằng sử dụng chatbot trong lĩnh vực dịch vụ khách hàng sẽ giúp doanh thu của các ngân hàng tăng lên khoảng 7,3 tỷ USD vào năm 2023 [19]. Chatbot được đánh giá là một công cụ hiệu quả để giảm thời gian phục vụ khách hàng, tăng tính tiện lợi cho khách hàng và giảm chi phí cho doanh nghiệp.

Chi phí mua một chatbot cũng phụ thuộc vào nhiều yếu tố khác nhau, bao gồm tính năng của chatbot, độ phức tạp của dự án, kích thước của doanh nghiệp, và quy mô triển khai, nhà cung cấp. Có thể kể đến Chatfuel - một nền tảng chatbot dựa trên Facebook Messenger, có gói miễn phí và các gói trả phí bắt đầu từ 15 USD mỗi tháng; hoặc Tars - một nền tảng chatbot giúp doanh nghiệp tạo ra các trang đích, các trang quảng cáo và các trang giới thiệu cho các sản phẩm hoặc dịch vụ của họ, có giá bắt đầu từ 99 USD mỗi tháng cho gói tiêu chuẩn.

Khách hàng sẽ gửi yêu cầu của mình thông qua khung chat và sẽ nhận được phản hồi từ chatbot giống như hình thức dịch vụ tương tác 1 trong phần 2.2.1. Mặc dù chatbot được cấu hình để giải quyết nhiều vấn đề khác nhau của khách hàng, từ đáp ứng các yêu cầu đơn giản đến hỗ trợ trong quy trình mua hàng hoặc dịch vụ sau bán hàng, tuy nhiên chatbot cũng có một số hạn chế như:

- Tính chính xác: Chatbot cần phải được lập trình chính xác để hiểu được các yêu cầu của khách hàng và cung cấp các câu trả lời phù hợp. Nếu không, chatbot có thể đưa ra câu trả lời sai hoặc không hiểu được yêu cầu của khách hàng.
- Tính nhạy cảm với ngôn ngữ: Chatbot có thể không hiểu được các yêu cầu phức tạp hoặc các câu trả lời phức tạp trong trường hợp khách hàng sử dụng ngôn ngữ không chính thức hoặc khó hiểu.

b. Hệ thống gợi ý sản phẩm

❖ Hệ thống gợi ý sản phẩm của Amazon:

Các hệ thống gợi ý đều được xây dựng phần lớn dựa trên các thuật toán máy học và cần thu thập thông tin liên quan đến sản phẩm, khách hàng và hành vi của họ để tiến hành nghiên cứu, đánh giá. Hiện nay, hệ thống gợi ý sản phẩm của các website TMĐT dần trở nên phổ biến và quan trọng. Chúng được sử dụng để giúp khách hàng tìm kiếm những sản phẩm họ có thể quan tâm dựa trên thông tin về thói quen mua sắm nhằm tăng hiệu suất bán hàng. Hệ thống này có hình thức dịch vụ thuộc loại 2 trong phần 2.2.1, hình thức mà hệ thống có thể tự phân tích nhu cầu của người dùng và đưa ra gợi ý cho người dùng.

Chúng tôi sử dụng website bán hàng Amazon cho cuộc khảo sát thực tế vì hệ thống gợi ý của Amazon được đánh giá là một trong những hệ thống thông minh và hiệu quả nhất hiện nay. Theo một số nghiên cứu, hệ thống gợi ý cải thiện doanh số bán hàng của Amazon lên 37% từ năm 2019 đến năm 2020. Theo nghiên cứu của McKinsey, doanh số bán hàng của công ty tăng trưởng lên tới 35% nhờ vào thuật toán đề xuất sản phẩm độc quyền [20].

Amazon sử dụng nhiều thuật toán khác nhau để cung cấp gợi ý sản phẩm cho khách hàng, trong đó thuật toán phổ biến nhất là Collaborative Filtering. Một số điểm đáng chú ý của hệ thống gợi ý sản phẩm của Amazon bao gồm:

- Khả năng đưa ra các đề xuất sản phẩm chính xác: Hệ thống gợi ý sản phẩm của Amazon sử dụng nhiều kỹ thuật khác nhau để phân tích dữ liệu khách hàng, đưa ra các đề xuất sản phẩm chính xác và phù hợp với sở thích của từng khách hàng.
- Tính tương tác: Hệ thống gợi ý sản phẩm của Amazon tích hợp tính năng tương tác, cho phép khách hàng đánh giá và đề xuất các sản phẩm họ quan tâm. Nhờ đó, hệ thống có thể học hỏi và cải thiện chất lượng đề xuất sản phẩm theo thời gian.

- Tính ứng dụng rộng: Hệ thống gợi ý sản phẩm của Amazon được sử dụng trong nhiều lĩnh vực, bao gồm các sản phẩm điện tử, đồ gia dụng, sách, phim ảnh, âm nhạc, thực phẩm và nhiều lĩnh vực khác.

Tuy nhiên, như mọi hệ thống, hệ thống gợi ý sản phẩm của Amazon cũng tồn tại một số hạn chế, ví dụ như:

- Phụ thuộc vào dữ liệu: Nếu dữ liệu không đủ lớn hoặc chất lượng dữ liệu không tốt, hệ thống sẽ không đưa ra các đề xuất sản phẩm chính xác.
- Không đảm bảo sự hài lòng của khách hàng: Mặc dù hệ thống gợi ý sản phẩm của Amazon đưa ra các đề xuất sản phẩm phù hợp với sở thích của khách hàng, tuy nhiên, điều này không đảm bảo rằng khách hàng sẽ hài lòng với sản phẩm mà họ mua.

Amazon không công bố chính xác khoản phí mà họ đã chi cho hệ thống gợi ý sản phẩm, tuy nhiên, để xây dựng được hệ thống gợi ý sản phẩm thành công như hiện tại, Amazon đã phải chi trả chi phí triển khai đắt đỏ thông qua các yếu tố sau:

- Phần cứng và mạng: Hệ thống gợi ý sản phẩm cần phải được triển khai trên nhiều máy chủ, mỗi máy chủ đều có khả năng xử lý và lưu trữ lớn.
- Phần mềm và công nghệ: Amazon sử dụng các công nghệ mới nhất để triển khai Hệ thống gợi ý sản phẩm của họ, điều này cần phải có các kỹ sư và chuyên gia về phần mềm và công nghệ để giúp triển khai hệ thống.
- Dữ liệu: Hệ thống gợi ý sản phẩm của Amazon yêu cầu lượng dữ liệu lớn để đưa ra các đề xuất sản phẩm chính xác và hiệu quả. Do đó, việc thu thập, lưu trữ và quản lý dữ liệu sẽ là một phần quan trọng trong việc triển khai hệ thống.

Không chỉ riêng Amazon mà còn đối với các trang TMĐT khác hiện có triển khai hệ thống gợi ý sản phẩm, chi phí triển khai hệ thống gợi ý sản phẩm là một con số khổng lồ, điều này gây áp lực lớn lên các trang TMĐT vừa và nhỏ.

3.1.2. Phân tích, kết luận

Thông qua việc khảo sát hệ thống thực tế, chúng tôi nhận ra không có hoặc có rất ít hệ thống gợi ý sản phẩm dựa trên dịch vụ tương tác. Hầu hết các hệ thống đều chỉ hỗ trợ 1 trong 2 yêu cầu của chúng tôi: Hệ thống dịch vụ tương tác hoặc hệ thống gợi ý sản phẩm. Điều này chứng tỏ hệ thống gợi ý sản phẩm dựa trên dịch vụ tương tác sẽ là đề tài nghiên cứu được mọi người quan tâm trong tương lai gần.

Ngoài ra, đối với các doanh nghiệp vừa và nhỏ, việc đầu tư vào cài đặt và quản lý hệ thống gợi ý là một việc tốn chi phí đáng kể. Những mô hình của những trang TMĐT lớn đã công bố đều nhằm phục vụ nhu cầu khổng lồ và đặc thù của trang lớn đó, doanh nghiệp vừa và nhỏ khó có thể tái sử dụng, nên nếu muốn đầu tư xây đòi hỏi doanh nghiệp nhỏ và vừa phải tự đầu tư nghiên cứu. Đối với một doanh nghiệp vừa và nhỏ với nhiều hạn chế về kinh phí, tài nguyên và nguồn lực, đây thật sự là một thách thức.

Nhằm hỗ trợ các doanh nghiệp vừa và nhỏ vượt qua thách thức trên, mô hình dịch vụ gợi ý tương tác trong luận văn này được đề xuất và xây dựng “hướng dịch vụ”. Theo mô hình này, các doanh nghiệp vừa và nhỏ chỉ cần cung cấp các thông tin cần thiết để dịch vụ gợi ý chúng tôi xây dựng có đủ dữ liệu triển khai gợi ý. Kết quả gợi ý sẽ được tích hợp vào website của các doanh nghiệp. Do đó, các doanh nghiệp không cần dành quá nhiều công sức và tài nguyên vào việc xây dựng, vận hành và quản lý một hệ thống gợi ý riêng biệt.

Một cách cụ thể, các doanh nghiệp thương mại điện tử, tập trung vào phân khúc doanh nghiệp vừa và nhỏ, sẽ nhúng một đoạn mã của chúng tôi vào website của họ, đoạn mã này đã được tích hợp các chức năng cần thiết liên quan đến việc gợi ý sản phẩm. Do sự phát triển nhanh chóng của internet và tầm quan trọng của cá nhân hóa người dùng, chúng tôi mong muốn cung cấp dịch vụ tương tác (interactive service) nhằm tăng trải nghiệm và sự hài lòng của người dùng. Mục

tiêu của chúng tôi là bắt các sự kiện (click, scroll, v.v) diễn ra trên website của doanh nghiệp, sau đó sử dụng các thuật toán để phân tích và đưa ra gợi ý các sản phẩm có liên quan.

3.2. Framework Trivi

Framework Trivi là mô hình được công bố trong bài báo Trivi: A Conceptual Framework for Customer Intelligence Systems for Small and Medium-sized Enterprises [6], được đề xuất với mục đích hỗ trợ và phục vụ cho các doanh nghiệp vừa và nhỏ trong việc phát triển và triển khai hệ thống thông minh về khách hàng (customer intelligence system). Framework Trivi giúp cho doanh nghiệp xác định và sử dụng được đúng dữ liệu và đúng kỹ thuật phân tích trong kỷ nguyên dữ liệu lớn ngày nay.

3.2.1. Nguyên tắc của hệ thống thông minh về khách hàng

Một hệ thống thông minh về khách hàng (customer intelligence system) của doanh nghiệp vừa và nhỏ bao gồm 4 tầng là tầng dữ liệu (data level), tầng thông tin (information level), tầng phân tích (analytics level) và tầng quyết định (decision level). Ngoài ra, còn có 1 tầng “data source” là tầng cung cấp dữ liệu cho hệ thống (thông thường tầng này sẽ nằm bên ngoài hệ thống)

- Tầng dữ liệu: bao gồm các dữ liệu từ nhiều nguồn và có thể phân thành các loại như dữ liệu nhân khẩu học, dữ liệu giao tác, dữ liệu hành vi.
- Tầng thông tin: tầng này đề cập đến các công nghệ mô tả, tổ chức và tích hợp các dữ liệu. Theo Trivi, dữ liệu khách hàng được biến đổi thành thông tin ở 4 dạng biết-cái gì (know-what), biết-ai (know-who), biết-như thế nào (know-how), biết-tại sao (know-why).
- Tầng phân tích: tập trung vào kỹ thuật phân tích như mô tả, dự đoán.
- Tầng quyết định: liên quan đến các công cụ hỗ trợ ra quyết định như báo cáo (dashboard), khuyến nghị (recommendations)

3.2.2. Mô hình dữ liệu

Để tổ chức dữ liệu từ nhiều nguồn khác nhau thành một cấu trúc phù hợp cho phân tích kinh doanh, một mô hình dữ liệu cần có bốn thành phần kiến thức như ở tầng thông tin biết-cái gì (know-what), biết-ai (know-who), biết-như thế nào (know-how), biết-tại sao (know-why).

- Biết ai (know-who): liên quan đến tất cả dữ liệu về các bên liên quan (stakeholders) của hệ thống
- Biết cái gì (know-what): gồm các dữ liệu về sản phẩm/dịch vụ mà nó có lợi cho việc xem, phân tích, đánh giá thông tin chi tiết về sản phẩm/dịch vụ cũng như đề xuất các sản phẩm/dịch vụ tiềm năng.
- Biết như thế nào (know-how): liên quan đến các dữ liệu giao dịch, tương tác của khách hàng mà nó có lợi cho việc xem, phân tích, đánh giá về khách hàng.
- Biết tại sao (know-why): liên quan đến các dữ liệu về sở thích và xếp hạng của khách hàng, có thể được sử dụng để đề xuất sản phẩm.

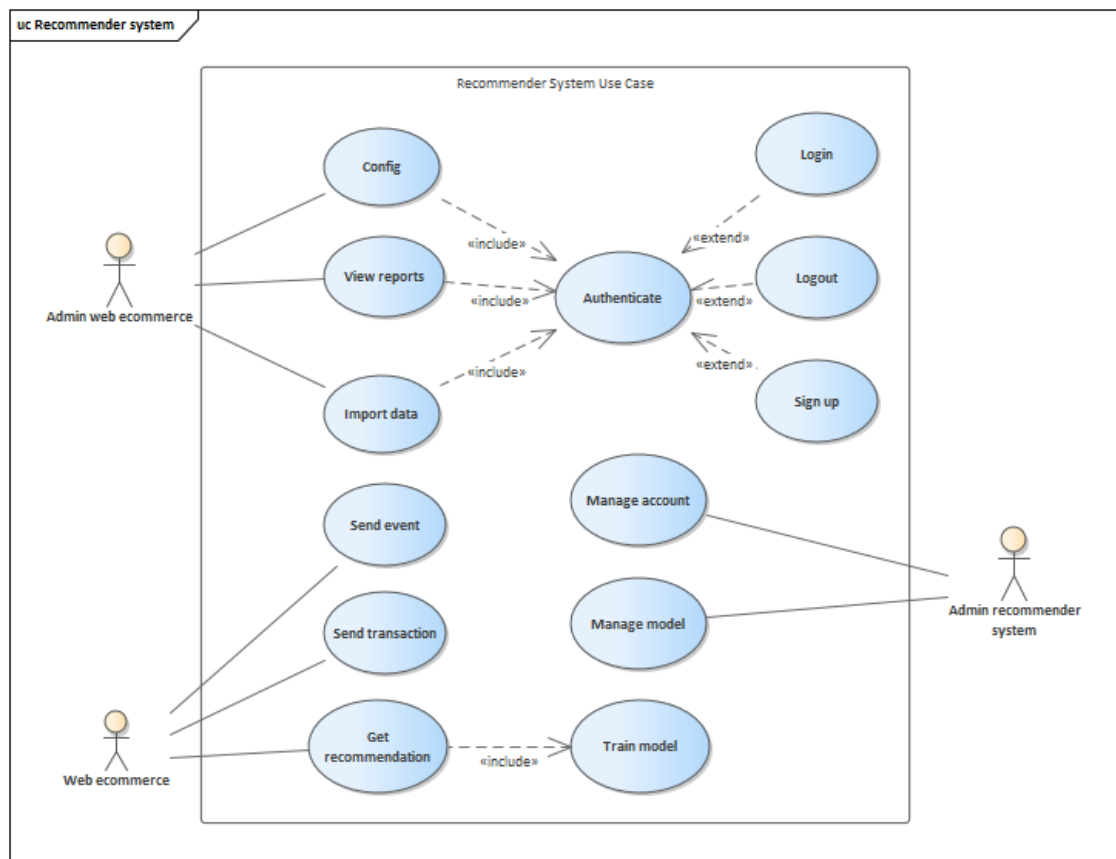
3.3. Thiết kế giải pháp

Dựa trên framework Trivi, trong phần này, chúng tôi thiết kế một giải pháp cho hệ thống gợi ý hướng dịch vụ bao gồm:

- Sơ đồ use case mô tả chức năng của hệ thống từ góc độ khách hàng doanh nghiệp.
- Mô hình dữ liệu kế thừa từ Framework Trivi [6] đã được chỉnh sửa và bổ sung cho phù hợp với dữ liệu của lĩnh vực thương mại điện tử ở mảng gợi ý sản phẩm.
- Kiến trúc hệ thống xác định các thành phần thiết yếu của một hệ thống kiến nghị hướng dịch vụ.

3.3.1. Sơ đồ use cases.

Mô hình dịch vụ gợi ý đặt ra các yêu cầu về việc quản lý tài khoản, lưu trữ dữ liệu, ghi nhận sự kiện xảy ra trên website thương mại điện tử và áp dụng thuật toán để gợi ý sản phẩm liên quan đến nhu cầu của người dùng nhất. Chúng tôi đã phân tích và mô hình hóa các chức năng dưới dạng sơ đồ Use Case như sau:



Hình 6. Sơ đồ Use Case Dịch vụ tư vấn sản phẩm

❖ Bảng 1 mô tả các tác nhân (actor) và vai trò trong hệ thống gợi ý sản phẩm:

| Tác nhân | Vai trò |
|---------------------------------|--|
| Admin website e-commerce | <p>Quản trị viên của khách hàng doanh nghiệp sẽ đăng ký website TMĐT của họ trên hệ thống dịch vụ, cung cấp những dữ liệu cần thiết và nhúng đoạn mã vào website TMĐT của họ.</p> <p>Admin web TMĐT thực hiện 4 chức năng chính:</p> <ul style="list-style-type: none"> - Đăng ký tài khoản vào trang quản lý của recommender system - Chọn chiến lược gợi ý cho từng trang của website TMĐT cũng như chu kỳ huấn luyện chiến lược. - Xem báo cáo thống kê - Nhập dữ liệu (thông tin sản phẩm, người dùng) |
| Website ecommerce | <p>Sau khi admin hoàn thành các thao tác đăng ký và cung cấp dữ liệu, website TMĐT lúc này đã có thể bắt và gửi các sự kiện về hệ thống dịch vụ và nhận kết quả gợi ý.</p> <p>Website TMĐT có 3 chức năng chính sau:</p> <ul style="list-style-type: none"> - Nhận biết và gửi sự kiện (click, scroll) đến hệ thống gợi ý sản phẩm. - Nhận biết và gửi giao tác (mua hàng) đến hệ thống gợi ý sản phẩm. - Nhận kết quả gợi ý từ hệ thống |
| Admin recommender system | Đây là quản trị viên của hệ thống gợi ý sản phẩm. |

| Tác nhân | Vai trò |
|----------|---|
| | Admin recommender system thực hiện 2 chức năng chính: - Tạo và quản lý tài khoản của hệ thống - Quản lý mô hình gợi ý sản phẩm, bao gồm việc kiểm tra, theo dõi tiến trình thực hiện của hệ thống |

Bảng 1. Mô tả tác nhân trong hệ thống

❖ Bảng 2 mô tả tổng quan các chức năng có trong hệ thống:

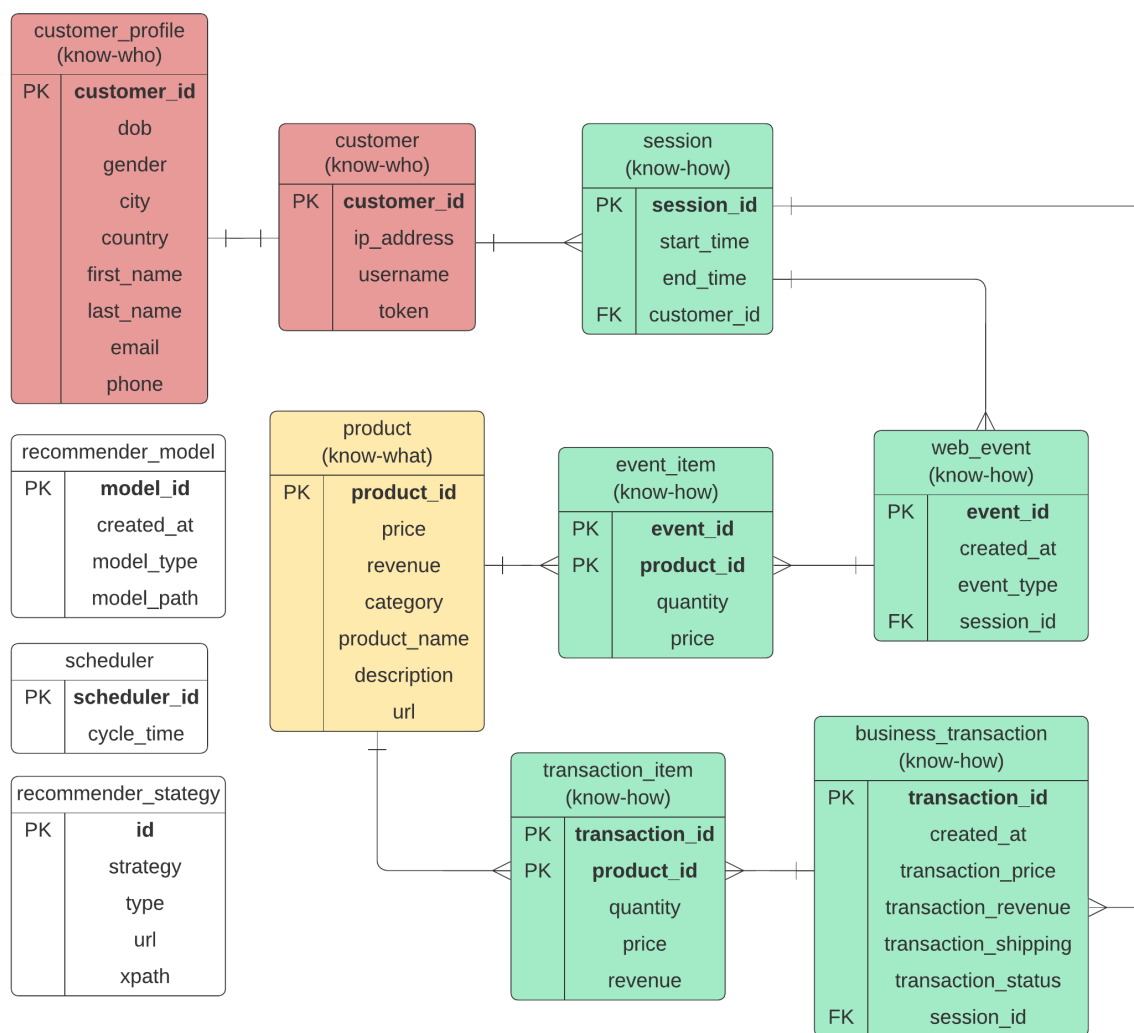
| ID | Tên | Mô tả |
|-------|--------------|---|
| UC1 | Authenticate | Hệ thống xác thực người dùng |
| UC1.1 | Login | Khách hàng doanh nghiệp đăng nhập vào hệ thống gợi ý sản phẩm. Khi login thành công, hệ thống sẽ tạo token để lưu tài khoản để đăng nhập nhanh cho những lần sau |
| UC1.2 | Sign up | Khách hàng doanh nghiệp đăng ký tài khoản vào hệ thống |
| UC1.3 | Logout | Khách hàng doanh nghiệp đăng xuất khỏi hệ thống |
| UC2 | View reports | Khách hàng doanh nghiệp xem thông tin liên quan đến báo cáo hàng tháng (danh sách sản phẩm phổ biến, thông tin các sự kiện diễn ra trên website) thông qua giao diện của dashboard. |
| UC3 | Import data | Khách hàng doanh nghiệp nhập dữ liệu vào hệ thống gợi ý sản phẩm, có thể bao gồm các thông |

| ID | Tên | Mô tả |
|------|--------------------|---|
| | | tin liên quan đến sản phẩm, khách hàng doanh nghiệp (được định dạng theo format quy định sẵn). |
| UC4 | Send event | Website TMĐT gửi dữ liệu bao gồm thao tác click, thêm vào giỏ hàng (add to cart), bỏ khỏi giỏ hàng (remove from cart) đến máy chủ gợi ý. |
| UC5 | Send transaction | Website TMĐT gửi giao tác bao gồm các thông tin liên quan đến việc mua bán đến máy chủ gợi ý. |
| UC6 | Get recommendation | Website TMĐT nhận phản hồi kết quả danh sách sản phẩm gợi ý từ máy chủ gợi ý. |
| UC7 | Train model | Khi nhận được lệnh gợi ý sản phẩm, website TMĐT nhận kết quả từ mô hình trí tuệ nhân tạo đã được huấn luyện theo lịch trình của quản trị viên khách hàng doanh nghiệp chọn. |
| UC8 | Config | Khách hàng doanh nghiệp chọn chiến lược gợi ý và thời gian chạy thuật toán trên giao diện. |
| UC9 | Manage account | Admin hệ thống gợi ý sản phẩm thêm, xóa, chỉnh sửa tài khoản trong hệ thống gợi ý sản phẩm. |
| UC10 | Manage model | Admin hệ thống gợi ý sản phẩm theo dõi tiến trình thực hiện của hệ thống gợi ý sản phẩm thông qua Airflow. |

Bảng 2. Đặc tả chức năng

3.3.2. Sơ đồ thực thể quan hệ (ERD)

Để lưu trữ dữ liệu đáp ứng các chức năng trên, đồng thời dựa trên mô hình cơ sở dữ liệu được công bố tại luận văn Trivi: A Conceptual Framework for Customer Intelligence Systems for Small and Medium-sized Enterprises [6], chúng tôi rút ra được cơ sở dữ liệu quan hệ như sau:



Hình 7. Lược đồ thực thể kết hợp – phần quản lý dữ liệu

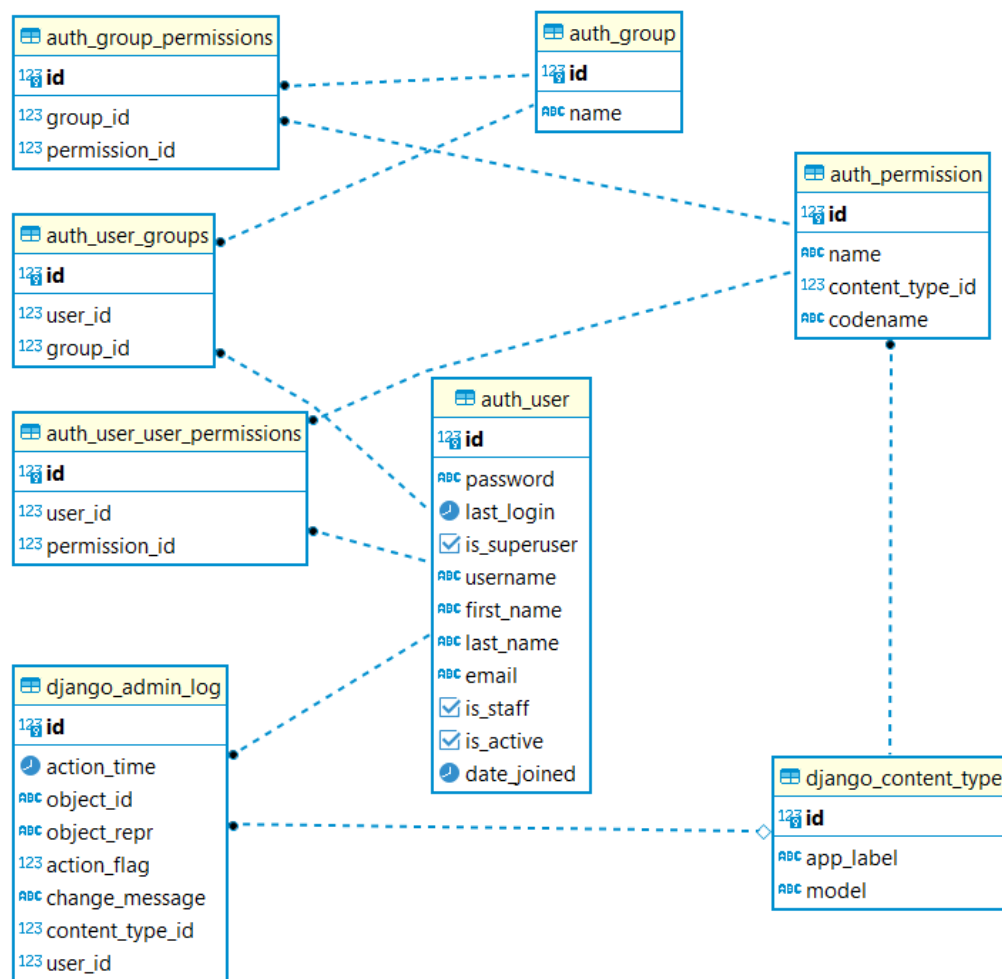
❖ Các thực thể được mô tả qua bảng 3:

| Thực thể | Mô tả |
|--|--|
| Customer (know-who) | Thông tin người dùng, bao gồm địa chỉ IP, tên đăng nhập và token đăng nhập vào web thương mại điện tử. |
| CustomerProfile (know-who) | Thông tin cá nhân của người dùng: tên, email, số điện thoại, giới tính, địa chỉ |
| Product (know-what) | Thông tin sản phẩm, bao gồm tên sản phẩm, giá, phân loại, mô tả và đường dẫn đến sản phẩm |
| Session (know-how) | Thông tin phiên đăng nhập của người dùng |
| BusinessTransaction (know-how) | Thông tin giao tác mua hàng của người dùng, cho biết phí, trạng thái, và thuộc phiên đăng nhập nào |
| TransactionItem (know-how) | Chi tiết lần mỗi lần mua hàng của người dùng, bao gồm định danh sản phẩm, số lượng, giá tiền và thuộc giao tác nào |
| WebEvent (know-how) | Thông tin sự kiện diễn ra trên web (click, scroll) |
| EventItem (know-how) | Chi tiết sự kiện thuộc về sản phẩm nào, bao gồm định danh sản phẩm, giá tiền và thuộc sự kiện nào |
| ManageAccount | Thông tin tài khoản admin. Mỗi account sẽ được gán cho một cơ sở dữ liệu để quản lý. |
| Scheduler | Quản lý thời gian và chiến lược chạy dịch vụ gợi ý sản phẩm. Chúng tôi lưu 3 giá trị: 1 – 1 day (1 ngày/1 lần chạy), 7 – 1 week (1 tuần/1 lần chạy), 30 – 1 month (1 tháng/1 lần chạy) |
| RecommendTrategy | Chiến lược gợi ý cài đặt theo nhu cầu. Ở đây có các chiến lược sau: Collaborative Filtering, Content-based, |

| Thực thể | Mô tả |
|----------|--|
| | Demographic, Hot items. Cần cung cấp XPath để có thể cài đặt chiến lược riêng cho từng trang. Ví dụ XPath có dạng: /html/body/div[1]/div[1]/main/div/a/img |

Bảng 3. Mô tả thực thể

Ngoài ra, chúng tôi minh họa riêng cơ sở dữ liệu phân quản lý tài khoản qua hình 5:



Hình 8. Sơ đồ thực thể quan hệ (Entity Relationship Diagram) – phân quản lý tài khoản

Chúng tôi hoàn toàn sử dụng module có sẵn (django.contrib.auth) của Django để quản lý tài khoản, vì vậy sơ đồ ER phần quản lý tài khoản sẽ theo mô hình quản lý tài khoản của Django.

3.3.3. Thiết kế kiến trúc hệ thống

Tiếp đến, chúng tôi xây dựng một kiến trúc hệ thống dịch vụ để phù hợp với các yêu cầu bên trên. Hệ thống dịch vụ gồm có 3 cụm chính, cụm cơ sở dữ liệu, cụm hệ thống gợi ý, cụm khách hàng doanh nghiệp được liên kết chặt chẽ với nhau. Liên kết các cụm này với 4 tầng trong Framework Trivi (phần 3.2) (tầng dữ liệu, tầng thông tin, tầng phân tích, tầng quyết định), cụm khách hàng doanh nghiệp tương ứng với tính chất tầng dữ liệu nguồn, cụm cơ sở dữ liệu tương ứng với tầng dữ liệu và tầng thông tin, cụm máy chủ gợi ý tương ứng với 2 tầng cấp cao là tầng phân tích và tầng quyết định. Ngoại trừ tầng dữ liệu nguồn ở cụm khách hàng doanh nghiệp, các tầng còn lại bao gồm cụm cơ sở dữ liệu và cụm hệ thống gợi ý có thể được xem là một hệ thống thông minh về khách hàng (customer intelligent system)

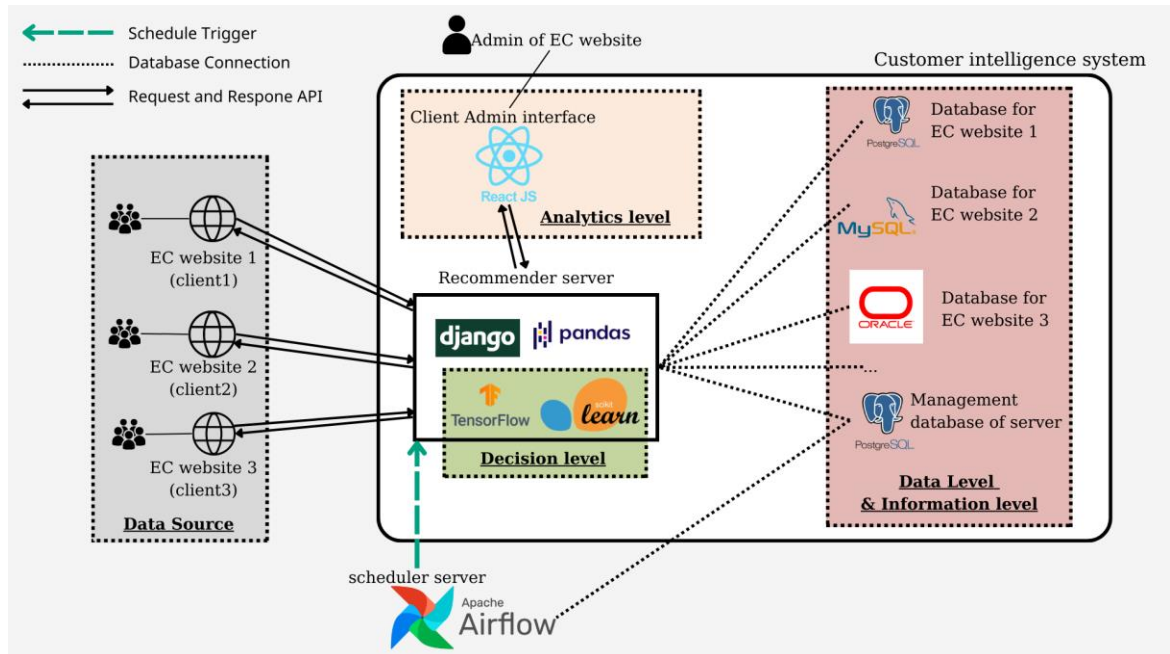
❖ Cụm khách hàng doanh nghiệp:

- Thành phần: bao gồm các trang website TMĐT. Ở cụm này, tất cả máy chủ đều thuộc quyền sở hữu và truy cập của khách hàng doanh nghiệp.
- Kết nối: Các website TMĐT này sẽ cần được nhúng một đoạn mã code javascript để có thể tương tác với hệ thống. Đoạn code này sẽ thực hiện việc bắt sự kiện và gửi sự kiện cho máy chủ gợi ý. Sau đó sẽ nhận lại phản hồi từ máy chủ.
- Đặc tính: Có đặc tính và tính chất của tầng dữ liệu nguồn (data source). Dữ liệu chủ yếu được lấy và thu thập từ cụm này.

❖ Cụm cơ sở dữ liệu:

- Thành phần: gồm các Cơ sở dữ liệu cung cấp cho các trang TMĐT theo mô hình thực thể quan hệ đã được trình bày ở phần 3.3.2, và có thể sử dụng bất kỳ cơ sở dữ liệu quan hệ.
 - Kết nối: các cơ sở dữ liệu được duy trì kết nối liên tục với recommender server để luôn trong trạng thái sẵn sàng bằng một vùng kết nối (connection pool) duy nhất.
 - Đặc tính: Có đặc tính và tính chất của tầng dữ liệu (data level) và tầng thông tin (information level). Cụm này sẽ giữ nhiệm vụ tổ chức và lưu trữ dữ liệu.
- ❖ **Cụm Hệ thống gợi ý:**
- Thành phần: gồm máy chủ gợi ý được xây dựng bằng framework Django, giao diện quản trị viên là một máy chủ cho khách hàng doanh nghiệp có thể đăng ký và tải dữ liệu cần thiết lên hệ thống được xây dựng bằng React JS và máy chủ lập lịch là một máy chủ quản lý các tác vụ được thực hiện theo thời gian biểu cố định được xây dựng bằng Airflow.
 - Kết nối: Máy chủ gợi ý là nơi cung cấp API cho toán bộ kiến trúc.
 - Đặc tính: Cụm này gồm có đặc tính của 2 tầng. Với tầng phân tích (analytics level) là đặc tính của giao diện khách hàng doanh nghiệp, vì ở đây có cung cấp cho khách hàng doanh nghiệp xem các dashboard phân tích. Với tầng quyết định (decision level) sẽ là đặc tính của các phần thuật toán gợi ý đã được cài đặt trong máy chủ gợi ý.

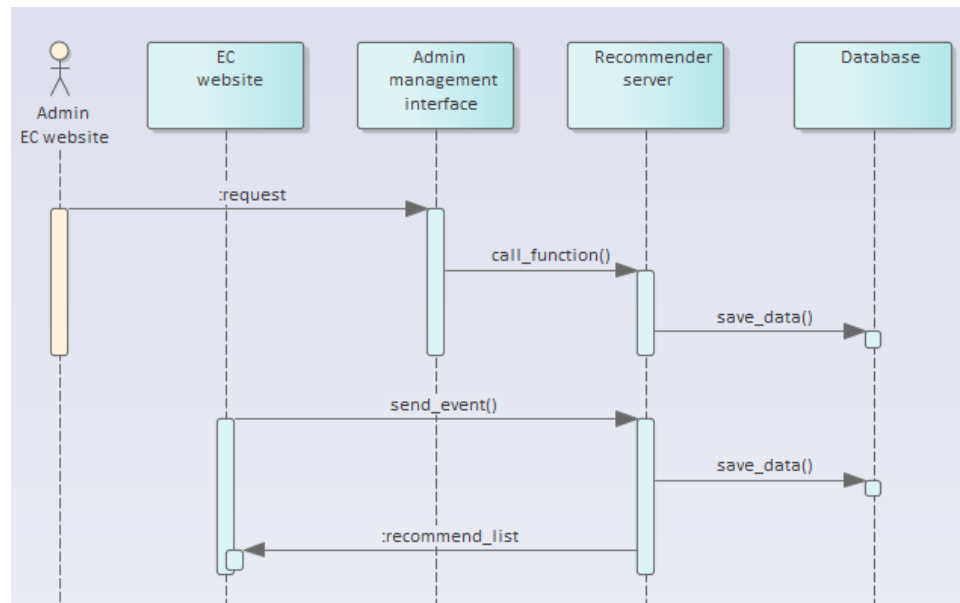
Kiến trúc hệ thống được minh họa như hình bên dưới:



Hình 9. Kiến trúc hệ thống

3.3.4. Sơ đồ Sequence

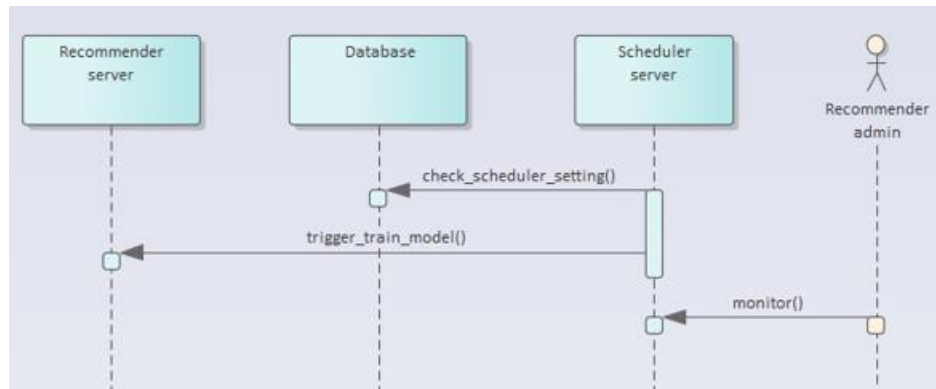
Dựa trên kiến trúc hệ thống mà chúng tôi đề xuất, sơ đồ tuần tự ở hình 9 và hình 10 mô tả tổng quát các giao tiếp giữa các thành phần trong hệ thống:



Hình 10. Sơ đồ tuần tự của hệ thống gợi ý

Thông qua giao diện quản lý của admin (admin management interface), admin website TMĐT (admin EC website) thực hiện các chức năng liên quan đến đăng ký, đăng nhập, xem báo cáo, nhập dữ liệu, cài đặt thông số cho hệ thống. Yêu cầu sẽ được gửi đến máy chủ gợi ý (recommender server) để xử lý, nếu thành công sẽ kết nối đến CSDL (database) để thực hiện lưu trữ.

Khi admin website TMĐT nhúng đoạn mã hệ thống vào website của mình, máy chủ gợi ý sẽ bắt được các sự kiện diễn ra trên website, sau đó, bằng các thuật toán gợi ý đã được cài đặt, máy chủ gợi ý sẽ phản hồi danh sách sản phẩm gợi ý đến website TMĐT.



Hình 11. Sơ đồ tuần tự bộ lập lịch của hệ thống

Nói về bộ lập lịch của hệ thống, máy chủ lập lịch (scheduler server) sẽ kiểm tra thông số mà admin website TMĐT đã cài đặt thông qua giao diện. Thông số này được lưu ở CSDL, bao gồm thời gian của chu kỳ huấn luyện mô hình, chiến lược (thuật toán) ứng với mỗi trang trên website. Khi đến đúng chu kỳ, máy chủ lập lịch sẽ yêu cầu máy chủ gợi ý thực hiện huấn luyện mô hình. Ngoài ra, quản trị viên của hệ thống gợi ý có thể giám sát tiến trình thực hiện thông qua Airflow và có thể kịp thời xử lý nếu có vấn đề phát sinh.

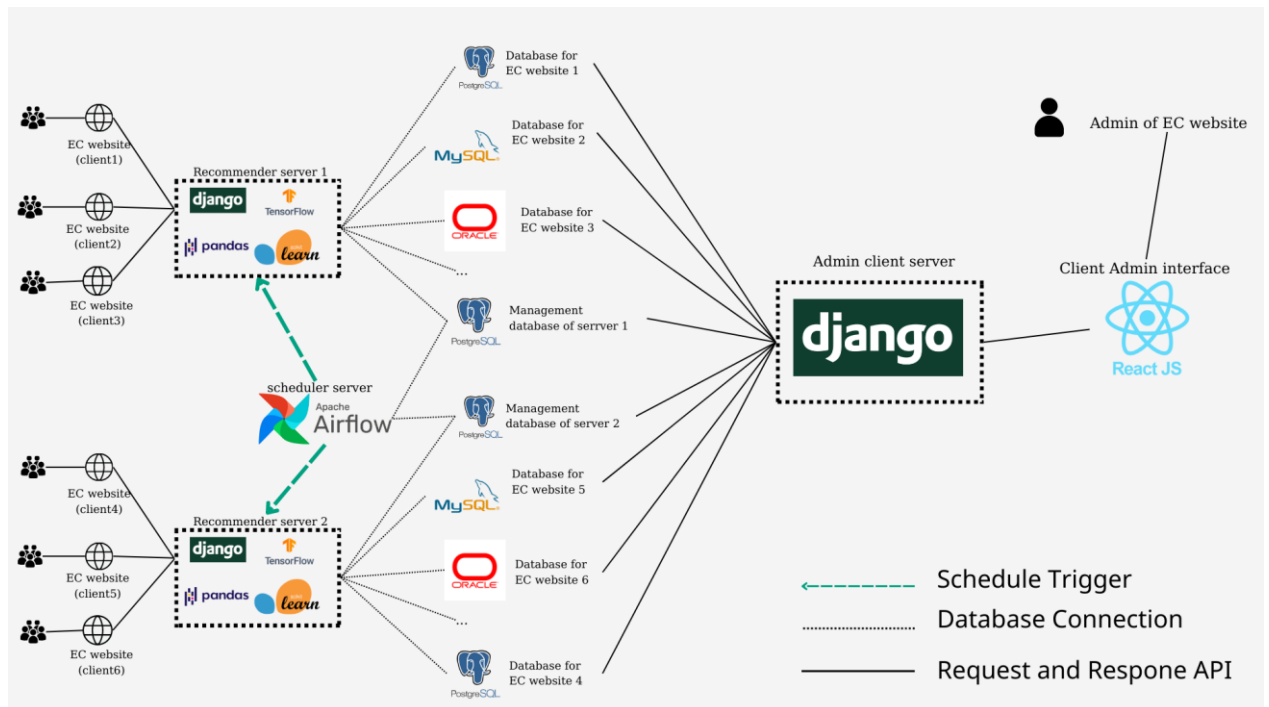
3.3.5. Đánh giá kiến trúc hệ thống

Với kiến trúc hệ thống ở hình 9, chúng tôi sẽ đánh giá kiến trúc hệ thống này qua một 4 tiêu chí là tính linh hoạt, tính dễ quản lý, tính mở rộng, độ phức tạp để có cái nhìn hoàn chỉnh hơn về hệ thống.

- ❖ Về tính linh hoạt, chúng tôi đánh giá theo 2 tiêu chí sau:
 - Linh hoạt thay đổi thành phần: các thành phần có thể thay đổi tùy ý bằng một công nghệ tương thích khác.
 - Linh hoạt trong việc phục vụ nhiều website TMĐT khác nhau: Mỗi trang TMĐT đều có các cấu trúc riêng, nhưng với hệ thống dịch vụ này, chúng tôi có thể phục vụ nhiều website TMĐT khác nhau cùng lúc dễ dàng.
- ❖ Tính dễ quản lý: Kiến trúc với các thành phần rõ ràng, số lượng thành phần phần nhỏ, được liên kết chặt chẽ với nhau theo cấu trúc cụ thể.
- ❖ Tính mở rộng: Đối với kiến trúc này, việc mở rộng để phục vụ cho 5 đến 10 website TMĐT rất dễ dàng, chỉ cần có đủ lượng cơ sở dữ liệu phục vụ cho mỗi trang. Tuy nhiên, nếu mở rộng để phục vụ cho 100 hay 1000 website TMĐT sẽ là một vấn đề nan giải. Vì với số lượng lớn trang TMĐT, recommender server sẽ gặp phải vấn đề quá tải. Lúc này, chúng tôi đề xuất 2 hướng giải quyết như sau:
 - Hướng giải quyết 1: Nhân bản hóa nhiều recommender server và tách biệt phần đăng ký và xem dashboard thành một máy chủ riêng lẻ và máy chủ bộ lập lịch sẽ chạy kích hoạt cho tất cả recommender server. Tuy nhiên, hướng giải quyết này có 1 điểm yếu là kiến trúc lúc này sẽ trở nên phức tạp hơn, và đồng thời nảy sinh vấn đề về phân phối khách hàng doanh nghiệp làm sao cho tối ưu hóa.
 - Hướng giải quyết 2: Thiết kế lại kiến trúc theo hướng microservice.Tuy nhiên, ở luận văn này, chúng tôi sẽ tập trung phân tích và phát triển kiến trúc như hình 9 và gác lại các phần mở rộng.

- ❖ Độ phức tạp: Kiến trúc hệ thống với độ khó vừa phải, số lượng thành phần ít, không cần quá nhiều nỗ lực và thời gian để hiểu rõ và áp dụng đúng cách và cải tiến.

Sau khi các thiết kế đã được chỉnh chu và rõ ràng, chúng tôi sẽ bắt đầu qua các bước cài đặt bên trong các thành phần của kiến trúc.



Hình 12. Kiến trúc hệ thống mở rộng

3.3. Cài đặt giải pháp

Ở phần này, chúng tôi sẽ làm rõ hơn các nhiệm vụ của từng thành phần tương ứng với kiến trúc đã thiết kế. Trọng tâm sẽ xoay quanh Recommender server (máy chủ gợi ý) với các nhiệm vụ cốt lõi của hệ thống được lập trình bằng Python. Bên cạnh đó, máy chủ bộ lập lịch sẽ đảm nhận các nhiệm vụ chạy theo thời gian sử dụng công cụ Airflow. Ngoài ra, chúng tôi còn nói về phương pháp phân phối dịch vụ cho khách hàng doanh nghiệp và cách bắt sự kiện người dùng.

3.3.1. Phương thức phân phối dịch vụ

Phương thức phân phối dịch vụ ở đây được hiểu là cách dịch vụ gợi ý của chúng tôi truyền và hiển thị kết quả gợi ý đến với trang website của doanh nghiệp. Trong luận văn này, chúng tôi đề xuất 2 phương thức như sau:

- ❖ Phương thức 1: Với mỗi lần chạy thuật toán thành công, chúng tôi sẽ lưu kết quả gợi ý dưới dạng JSON và gửi đến website của khách hàng doanh nghiệp. Đội ngũ phát triển của doanh nghiệp đó sẽ tiến hành xử lý kết quả và cài đặt hiển thị gợi ý trên website theo ý muốn của doanh nghiệp.

Phương thức này có những ưu điểm sau đây:

- Không can thiệp vào việc hiển thị của website TMĐT: Hệ thống của chúng tôi chỉ có nhiệm vụ trả kết quả là danh sách sản phẩm gợi ý mà không can thiệp vào giao diện của website TMĐT. Chúng tôi cho rằng, việc can thiệp đến giao diện của website TMĐT của doanh nghiệp có thể làm ảnh hưởng đến thiết kế và trải nghiệm người dùng của website đó.
- Tiện lợi cho việc xử lý: JSON là một định dạng dữ liệu đơn giản và phổ biến, có tốc độ truyền tải và xử lý nhanh. Ngoài ra, JSON cung cấp các phương thức đơn giản để thao tác với dữ liệu, cho phép người lập trình dễ dàng xử lý và truy xuất dữ liệu.

Tuy nhiên, phương thức này có một số điểm hạn chế cần quan tâm:

- Không hỗ trợ các tính năng phức tạp: Do đặc tính đơn giản, JSON không hỗ trợ các tính năng phức tạp như mã hóa, chữ ký số, điều này có thể làm giảm tính bảo mật của hệ thống.
 - Vấn đề về bảo mật: Như đã nói ở ý trên, việc gửi dữ liệu JSON đòi hỏi các biện pháp bảo mật chặt chẽ để đảm bảo an toàn cho dữ liệu.
- ❖ Phương thức 2: Với mỗi lần chạy thuật toán thành công, chúng tôi xử lý và hiển thị kết quả dưới dạng pop-up lên giao diện web của doanh nghiệp. Mỗi

khi người dùng tương tác với web, pop-up sẽ hiện lên với thông tin về sản phẩm được gợi ý và sẽ tự động đóng lại sau vài giây.

Phương thức này sẽ can thiệp vào giao diện người dùng, và có những ưu điểm sau đây:

- Cung cấp dịch vụ trọn vẹn: Phía doanh nghiệp không cần thực hiện bất kỳ công việc gì liên quan đến việc gợi ý sản phẩm. Chúng tôi đảm bảo cung cấp dịch vụ từ việc nhận và xử lý dữ liệu đến việc thiết kế, hiển thị cho người dùng của website TMĐT.
- Tăng tính tương tác: Bằng cách can thiệp vào giao diện của website của doanh nghiệp, có thể tạo ra các tính năng mới hoặc cải thiện tính năng hiện có để tăng tính tương tác của người dùng với website.

Dù việc can thiệp vào giao diện của website có thể mang lại một số lợi ích, tuy nhiên nó cũng tiềm ẩn nhiều nhược điểm và rủi ro tiềm tàng:

- Làm ảnh hưởng đến trải nghiệm người dùng: Khi can thiệp vào giao diện của website của đối tác, có thể làm thay đổi hoặc ảnh hưởng đến cấu trúc, thiết kế và trải nghiệm người dùng của website đó.
- Không an toàn: Việc can thiệp vào giao diện của website của đối tác có thể dẫn đến các lỗ hổng bảo mật, như làm giảm tính bảo mật của website đó hoặc tạo ra điểm tiếp xúc cho các kẻ tấn công tìm kiếm lỗ hổng bảo mật.
- Vi phạm quyền sở hữu trí tuệ: Giao diện của website là một phần của quyền sở hữu trí tuệ của doanh nghiệp, việc can thiệp vào giao diện này có thể vi phạm quyền sở hữu trí tuệ của doanh nghiệp.
- Về mặt hiển thị: Thiết kế của pop-up với thiết kế giao diện của website có thể không tương thích với nhau.

Theo nhận định của chúng tôi, phương thức 1 mang tính hiệu quả cao hơn khi doanh nghiệp có thể linh hoạt tùy chỉnh, cài đặt cách hiển thị kết quả lên giao diện website, đảm bảo sự tương thích với hệ thống của website TMĐT. Trong luận

văn này, chúng tôi sẽ minh họa hai phương thức trên và sẽ trình bày cụ thể ở chương 4.

3.3.2. Máy chủ gợi ý (Recommender server)

Máy chủ gợi ý là trung tâm của cả kiến trúc của chúng tôi, được lập trình bằng ngôn ngữ Python. Máy chủ gợi ý gồm 3 nhóm nhiệm vụ lớn:

- Cung cấp RESTful API

Recommender server sẽ tương tác với website ecommerce và giao diện admin bằng các API được cung cấp được lập trình bằng Django. Chúng tôi lập trình các thuật toán gợi ý theo như kiến thức đã tìm hiểu, hệ thống sẽ gửi request đến RESTful Route đã được khai báo để thực thi các chức năng tương ứng. Ngoài ra, chúng tôi cũng sử dụng các modules có sẵn của Django để một số việc quản lý như, xác thực người dùng, đăng ký, upload tập tin trở nên đơn giản và nhanh chóng hơn.

Một số modules của Django mà chúng tôi đã áp dụng:

- django.db: module cho phép tương tác với cơ sở liệu.
- django.urls: module cho phép khai báo url cho hệ thống.
- django.contrib.auth: module cung cấp tính năng đăng nhập, đăng ký và quản lý người dùng.
- django.http: cung cấp class và hàm liên quan đến việc xử lý HTTP requests và HTTP response, trả về các đối tượng HttpResponse và JsonResponse để trả lời HTTP requests.
- Kết nối với các database
 - Về kết nối

Recommender server sẽ luôn giữ với mỗi database một kết nối (connection pool). Thông qua bộ ORM do Django cung cấp, Recommender server có thể kết nối đến nhiều bộ cơ sở dữ liệu quan hệ và truy vấn một cách dễ dàng.

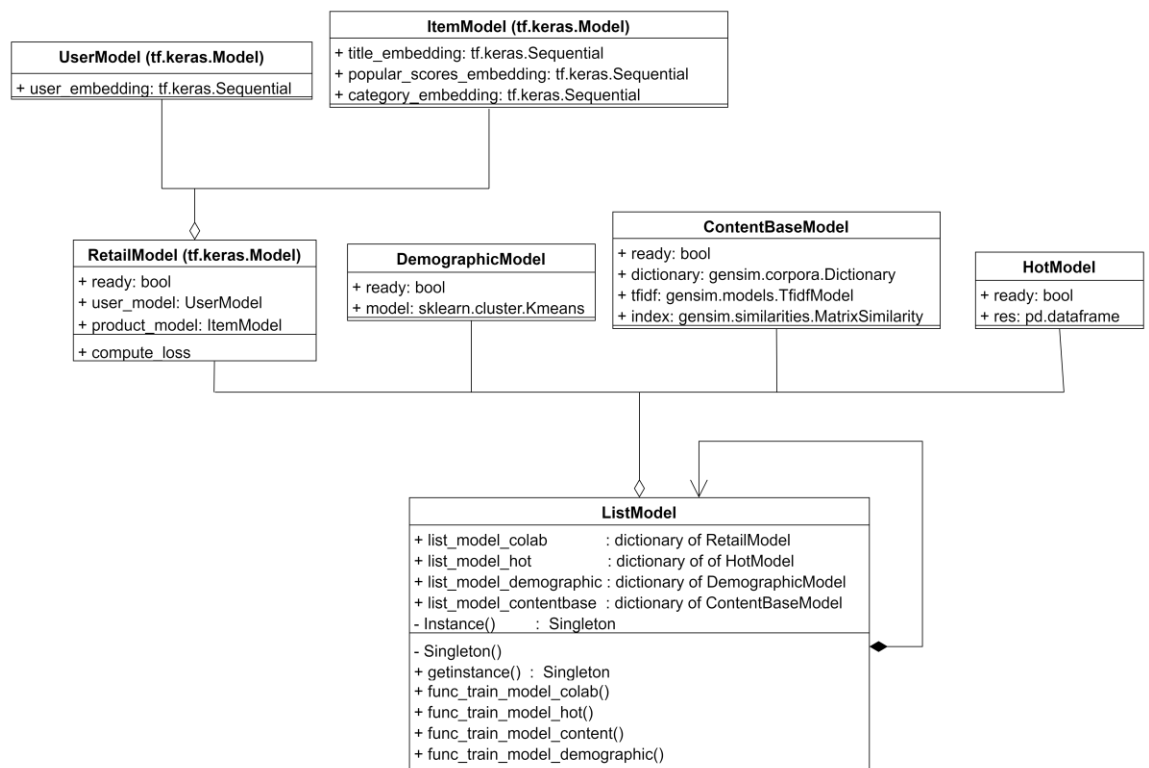
- Về truy vấn

Với mỗi lần truy vấn bằng bộ ORM do Django cung cấp, kết quả sẽ trả về dưới dạng QuerySet, tuy nhiên, ở dạng này chúng tôi rất khó sử dụng và sẽ chuyển sang dạng dataframe của Pandas để dễ sử dụng hơn.

- Các thuật toán gợi ý

Mỗi website TMDT đều có thể sử dụng 4 mô hình gợi ý khác nhau nên chúng tôi thực hiện như sau:

- Với mỗi mô hình thuật toán, chúng tôi đóng gói trong một lớp.
- Chúng tôi có xây dựng thêm một lớp Singleton để có tất cả các mô hình đó. Mục đích của Singleton là để các mô hình thuật toán đó được tạo ra duy nhất trong toàn bộ chương trình. Điều này sẽ giúp các mô hình thuật toán không phải nạp nhiều lần, giúp tối ưu thời gian và tài nguyên xử lý.



Hình 13. Recommender mẫu thiết kế hướng đối tượng

3.3.3. Máy chủ lập lịch (Scheduler server)

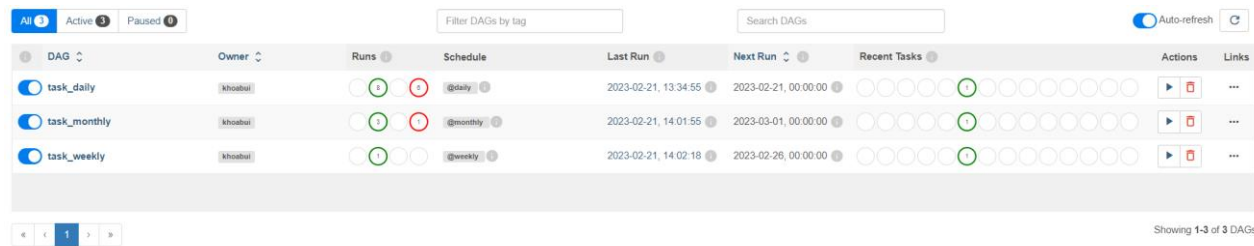
Scheduler server gồm có 3 nhiệm vụ chính được thiết lập sẵn bằng công cụ Airflow, các nhiệm vụ này giúp cho các mô hình thuật toán của website TMĐT được huấn luyện đúng thời gian mà khách hàng doanh nghiệp đã đăng ký. Ngoài ra, thuật toán sản phẩm phổ biến nhất chúng tôi sẽ cho huấn luyện mỗi ngày, thuật toán demographic và contentbase được huấn luyện ngay khi import dữ liệu:

- Huấn luyện các mô hình thuật toán theo ngày
- Huấn luyện các mô hình thuật toán theo tuần
- Huấn luyện các thuật toán theo tháng

Các nhiệm vụ đó sẽ được thực hiện với 3 bước:

- Bước 1: Truy vấn dữ liệu từ các cơ sở dữ liệu
- Bước 2: Gọi API vào recommender system để có thể huấn luyện thuật toán
- Bước 3: Sau khi nhận được kết quả trả về từ recommender system, sẽ thông báo nếu như quá trình huấn luyện có vấn đề.

Ngoài ra, sau khi cài đặt các nhiệm vụ trên, Airflow có cung cấp sẵn công cụ để quản lý các tác vụ, điều này giúp cho quản trị viên dễ dàng phát hiện vấn đề một cách nhanh chóng và kịp thời.



| DAG | Owner | Runs | Schedule | Last Run | Next Run | Recent Tasks | Actions | Links |
|--------------|---------|-------------------------------|----------|----------------------|----------------------|--|--|-------|
| task_daily | khoabui | 1 2 | @daily | 2023-02-21, 13:34:55 | 2023-02-21, 00:00:00 | 1 2 3 4 5 6 7 8 9 10 | ▶ ⏏ ⋮ | |
| task_monthly | khoabui | 1 2 | @monthly | 2023-02-21, 14:01:55 | 2023-03-01, 00:00:00 | 1 2 3 4 5 6 7 8 9 10 | ▶ ⏏ ⋮ | |
| task_weekly | khoabui | 1 2 | @weekly | 2023-02-21, 14:02:18 | 2023-02-26, 00:00:00 | 1 2 3 4 5 6 7 8 9 10 | ▶ ⏏ ⋮ | |

Hình 14. Màn hình làm việc của Airflow

3.3.4. Bắt sự kiện hành vi người dùng

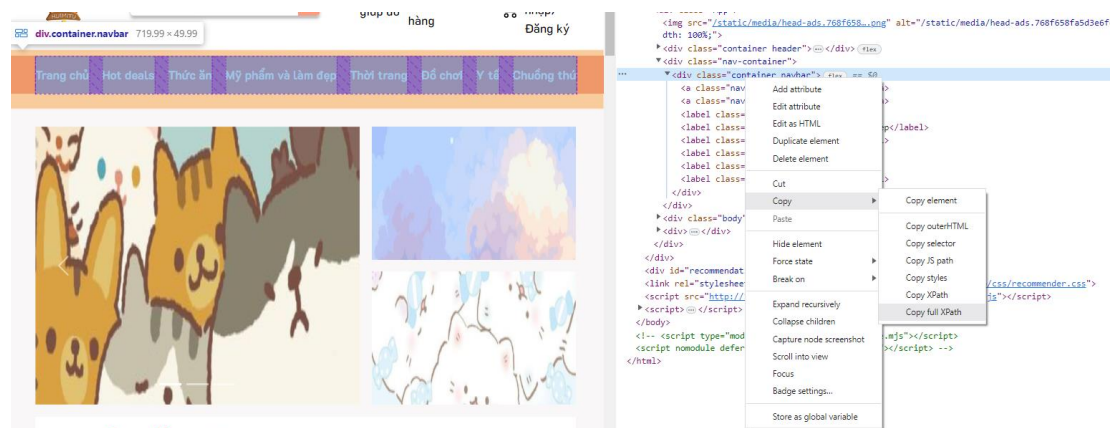
Như đã phân tích ở phần 2.2.1, việc nhận biết hành vi người dùng diễn ra trên website là vô cùng cần thiết và mang lại nhiều lợi ích. Tuy nhiên, để một hệ thống dịch vụ có thể bắt được sự kiện hành vi người dùng càng là một thách thức

khó khăn cần phải giải quyết. Ở phần này, chúng tôi sẽ đề xuất giải pháp cài đặt dịch vụ có thể bắt những sự kiện này qua 3 phần trình bày: dữ liệu khách hàng doanh nghiệp cung cấp, đoạn mã nhúng và cách thức hoạt động.

- Dữ liệu khách hàng doanh nghiệp cung cấp:

Mỗi website TMĐT khác nhau có thể có những cấu trúc khác nhau. Tuy nhiên, đối với các website TMĐT vừa và nhỏ, thông thường, những sản phẩm hiển thị trên cùng một trang (screen) đều tái sử dụng cùng thành phần (components), do đó, chúng có cùng đường dẫn (XPath) trên cấu trúc trang HTML đó.

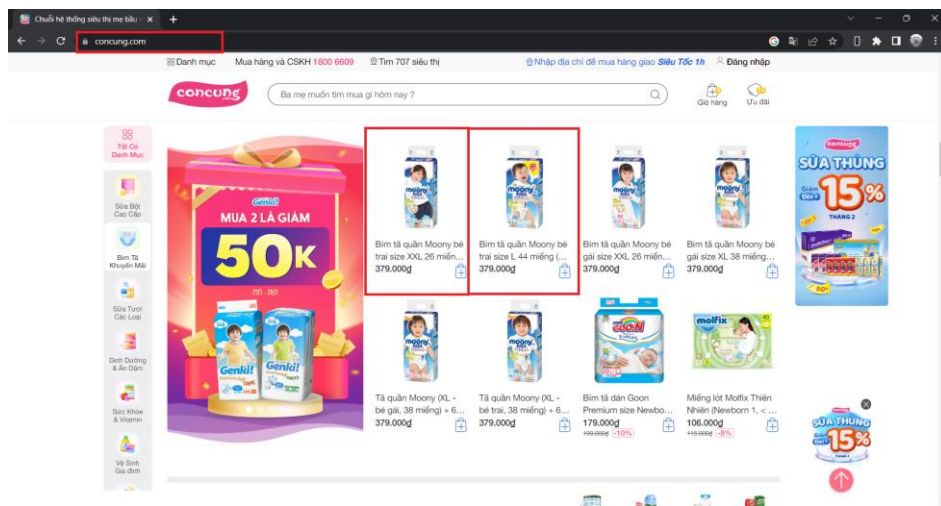
Từ ý tưởng đó, chúng tôi cần khách hàng doanh nghiệp (ở đây là người quản trị website TMĐT) cung cấp XPath và URL của các sản phẩm theo từng trang vì mỗi trang có XPath và URL khác nhau. Khách hàng doanh nghiệp có thể thực hiện sao chép XPath và URL một cách dễ dàng nhờ vào tính năng tiên tiến của trình duyệt web ngày nay.



Hình 15. Cách copy XPath bằng trình duyệt

Ví dụ, trên website của Concung, một sản phẩm bất kỳ ở homepage đều có chung cấu trúc XPath là “/html/body/div[5]/div[1]/section/div[2]/div[1]/div/div[1]/a” và URL là “https://concung.com/” để biểu thị trang chủ đang xét.

Tại đây xuất hiện một vấn đề cần được xem xét: Không phải bất kỳ XPath giống nhau (nhận định chúng được tạo nên bởi các components giống nhau) đều có cùng URL. Các URL này không giống nhau hoàn toàn nhưng sẽ có cùng định dạng. Lấy trang chi tiết sản phẩm của Concung làm ví dụ minh họa, URL sẽ có dạng “https://concung.com/{loai_hang}/{ten_san_pham}”, điều quan trọng là doanh nghiệp cần nhận biết được định dạng và thay đổi chúng với chuỗi bất kì thành “{string}”, chuỗi số thành “{number}”.

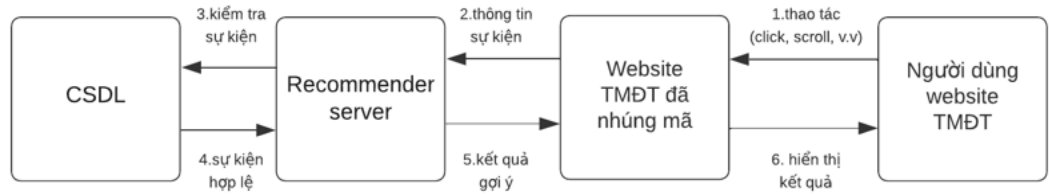


Hình 16. Trang chủ website Concung

- Đoạn mã nhúng:

Sau khi hệ thống đã có được dữ liệu, để website có thể giao tiếp với hệ thống, khách hàng doanh nghiệp cần nhúng một đoạn mã nhỏ vào đoạn mã thực thi của website TMĐT. Đoạn mã đó sẽ có 2 chức năng như sau: lắng nghe sự kiện trên website để biết được những XPath người dùng ấn và sau đó hiển thị sản phẩm gợi ý theo thuật toán đã được cài đặt.

- Cách thức hoạt động được minh họa tổng quát qua hình 17 như sau:



Hình 17. Cách thức bắt sự kiện trên website TMĐT

Sau khi người dùng ấn vào website TMĐT, đoạn mã đã nhúng sẽ lắng nghe thao tác của người dùng, xác định XPath (biểu thị element) tại URL (đường dẫn trang) mà người dùng vừa thao tác, sau đó gửi lên cho Máy chủ gợi ý (Recommender server). Recommender server sẽ lấy thông tin từ CSDL để kiểm tra thành phần mà người dùng đang tương tác đã cài đặt thuật toán gợi ý hay chưa, nghĩa là những hành động click lung tung trên website TMĐT sẽ không được xử lý. Nếu thỏa điều kiện, Recommender server sẽ phản hồi lại sự kiện đó một kết quả gợi ý.

Với cách làm này, website TMĐT của khách hàng doanh nghiệp sẽ không bị ảnh hưởng nhiều về hiệu năng vì Máy chủ gợi ý (Recommender server) ý hoạt động song song với website TMĐT của khách hàng doanh nghiệp, không ảnh hưởng đến các xử lý dữ liệu của doanh nghiệp trên website TMĐT.

3.4. Kết luận chương

Chương 3 trình bày mô hình triển khai dịch vụ gợi ý của chúng tôi dựa trên Framework Trivi. Các phân tích thiết kế và cài đặt ở chương này sẽ được chúng tôi thực thi hoàn chỉnh để chứng minh tính khả thi của luận văn. Kiến trúc được đề xuất tuy chưa thể đáp ứng hoàn toàn nhu cầu một dịch vụ ở ngoài thực tế, tuy nhiên nó là một nền tảng vững chắc để có thể cải thiện và nâng cấp ở tương lai.

CHƯƠNG 4: XÂY DỰNG HỆ THỐNG VÀ THỰC NGHIỆM

Chương 4 nhằm minh họa các bước xây dựng hệ thống và tiến hành thực nghiệm để minh họa kết quả. Trong chương này, vận dụng các kiến thức và mô hình đã trình bày ở các chương trước, chúng tôi thiết lập và cấu hình hệ thống, tổ chức các thành phần hệ thống và lập trình các thành phần phục vụ luận văn. Sau cùng, chúng tôi tiến hành thực nghiệm trên website TMDT mẫu để kiểm chứng kết quả.

4.1. Cài đặt package sử dụng

Phần này liệt kê các packages cần thiết để xây dựng hệ thống, các packages này là cơ sở để sử dụng các công cụ mà chúng tôi đã trình bày ở phần 2.4.

Có hai giai đoạn trong quá trình chuẩn bị trước khi triển khai các chức năng:

- Giai đoạn cài đặt: Cài đặt các packages cần thiết.
- Giai đoạn cấu hình: Cấu hình môi trường ảo, khởi động máy chủ, thiết lập kết nối giữa máy chủ gợi ý, máy chủ lập lịch và máy chủ cơ sở dữ liệu.

Phần này xin được liệt kê tất cả packages cần thiết cho luận văn này.

| Package | Mô tả | Phiên bản |
|---------------|---|-----------|
| Python | Hệ thống gợi ý sản phẩm được xây dựng trên Django, là một framework của Python | 3.7.9 |
| pip | Trình quản lý cho phép cài đặt các thư viện và module Python một cách dễ dàng | 20.1.1 |
| NodeJS | Giao diện quản lý dữ liệu của hệ thống được xây dựng trên ReactJS, là một framework front-end của | 14.17.3 |

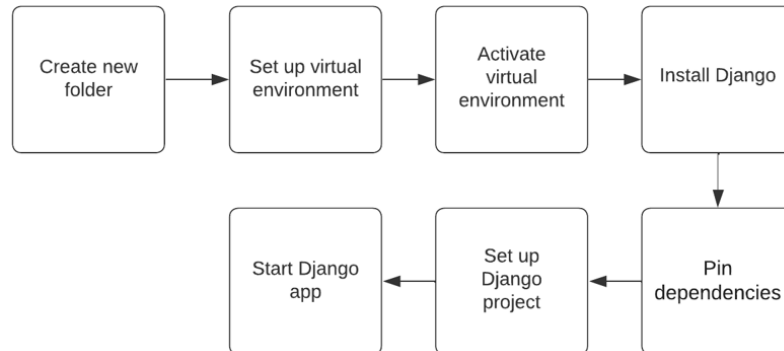
| Package | Mô tả | Phiên bản |
|-------------------|---|-----------|
| | Javascript. Để thực thi Javascript, cần cài đặt NodeJS | |
| npm | Node Package Manager (npm) là trình quản lý gói cho phép cài đặt, quản lý và chia sẻ các gói mã nguồn JavaScript. | 8.1.3 |
| virtualenv | Là công cụ tạo môi trường ảo Python độc lập, giúp tạo các folders cần thiết để thực thi đồ án | 20.16.6 |
| PostgreSQL | Hệ quản trị cơ sở dữ liệu quan hệ có mã nguồn mở, giúp quản lý dữ liệu | 14.4 |
| MySQL | Hệ quản trị cơ sở dữ liệu quan hệ có mã nguồn mở, giúp quản lý dữ liệu | 8.0.30 |
| Airflow | Bộ lập lịch có mã nguồn mở, giúp quản lý các nhiệm vụ chạy theo lịch | 2.4.2 |

Bảng 4. Các packages cần thiết trong hệ thống

4.2. Thiết lập và cấu hình

Sau khi hoàn tất việc cài đặt các packages, tiến hành set up database, kết nối hệ thống gọi ý với giao diện và database. Luận văn của chúng tôi phân ra thành hai phần chính: Hệ thống gọi ý sản phẩm (Django) và Giao diện quản lý (ReactJS).

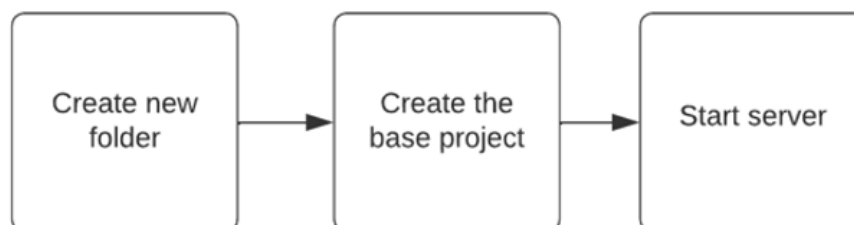
Về phía hệ thống gợi ý, cần tạo dự án trong môi trường Django thông qua quy trình như hình 14:



Hình 18. Quy trình tạo dự án Django

Trước tiên, tạo thư mục để lưu dự án, tiến hành set up và kích hoạt môi trường ảo. Lúc này, đường dẫn thư mục sẽ chuyển vào môi trường ảo, tại đây có thể cài đặt Django thông qua lệnh pip mà Python cung cấp. Khi cài đặt thành công Django, có thể lưu các phần phụ thuộc (dependencies) vào file text để có thể theo dõi phiên bản cài đặt Django đang quản lý. Việc lưu lại các dependencies còn hỗ trợ cho trường hợp tái sử dụng dự án, khi đó, chỉ cần thực thi lệnh cài đặt dependencies trong file text sẽ dễ dàng cài đặt đúng phiên bản Django của dự án. Cuối cùng thực thi lệnh set up dự án Django và khởi động Django app chứa chức năng cụ thể của hệ thống.

Về phía giao diện quản lý của admin, các bước tạo dự án Reactjs được mô tả qua hình 15:



Hình 19. Quy trình tạo dự án ReactJS

Trước hết, tạo thư mục mới để lưu trữ dự án, sau đó sử dụng công cụ npx được bao gồm trong npm để chạy lệnh thực thi tạo dự án ReactJS. Lệnh này sẽ cài đặt dự án ReactJS vào thư mục đã chỉ định, đồng thời chạy tải xuống bất kỳ dependencies bổ sung. Khi tạo thành công, màn hình sẽ hiển thị danh sách các lệnh npm cho phép xây dựng (build), chạy dự án (run), kiểm thử (test), bắt đầu sử dụng (start). Cuối cùng, khởi tạo máy chủ cục bộ và chạy dự án trong trình duyệt của mình.

Để kết nối CSDL, đầu tiên, cấu hình CSDL như sau:

```
DATABASES = {  
  'default': {  
    'ENGINE': 'django.db.backends.postgresql',  
    'NAME': env('DB_NAME'),  
    'USER': env('DB_USER'),  
    'PASSWORD': env('DB_PASSWORD'),  
    'HOST': env('DB_HOST'),  
    'PORT': env('DB_PORT'),  
    'OPTIONS': {"options": "project="+env('DB_ENDPOINT'),  
               "sslmode": "require"},  
  },  
}
```

Hình 20. Cấu hình CSDL

Hình 19 thể hiện việc cung cấp thông tin của CSDL, bao gồm: engine (tên của trình kết nối CSDL), tên CSDL, tên và mật khẩu tài khoản kết nối, cổng kết nối. Ngoài ra, nếu dùng cơ sở dữ liệu ở Serverless Postgres, chúng tôi cần cung cấp thêm các phần trong mục “options” gồm SSL và Endpoint để có thể truy cập cơ sở dữ liệu. Sau đó, kết nối các bảng ở dự án Django với CSDL thông qua lệnh migrate, trong đó, các bảng ở dự án Django theo mô hình đã được trình bày ở phần 3.2.2.

Để tạo một bảng (class) trong dự án Django, có thể thực hiện như sau:

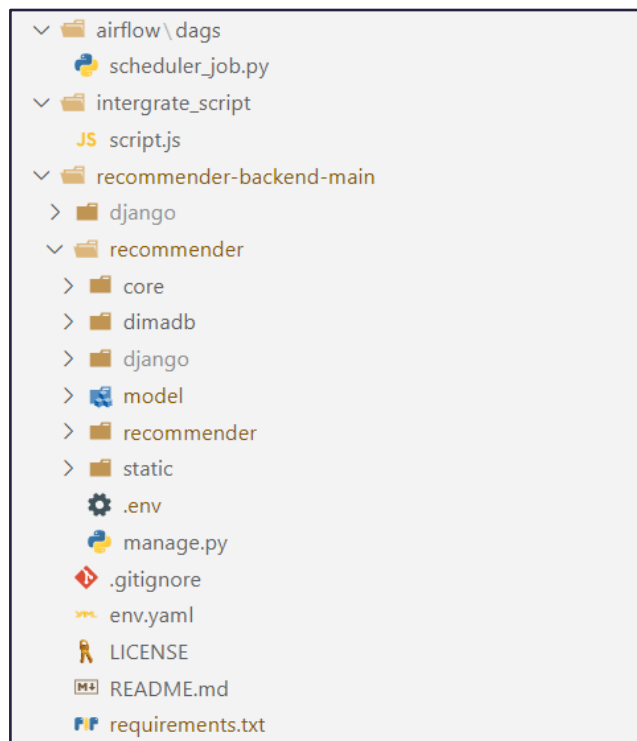
```
class Customer(models.Model):
    customer_id = models.CharField(primary_key=True, max_length=50, null=False, blank=True)
    username = models.CharField(max_length=30, null=False, blank=True)
    token = models.CharField(max_length=100, null=True, blank=True)
```

Hình 21. Minh họa tạo bảng trong dự án Django

4.3. Cấu trúc thư mục mã nguồn hệ thống

Sau khi thiết lập và cấu hình hệ thống, chúng tôi xin trình bày kiến trúc các thành phần hệ thống, bao gồm hệ thống gợi ý sản phẩm và giao diện khách hàng doanh nghiệp. Phần này giải thích các chức năng của từng thư mục để có thể dễ dàng thêm, chỉnh sửa các module hệ thống khi có vấn đề mới phát sinh.

4.3.1. Thành phần gợi ý sản phẩm



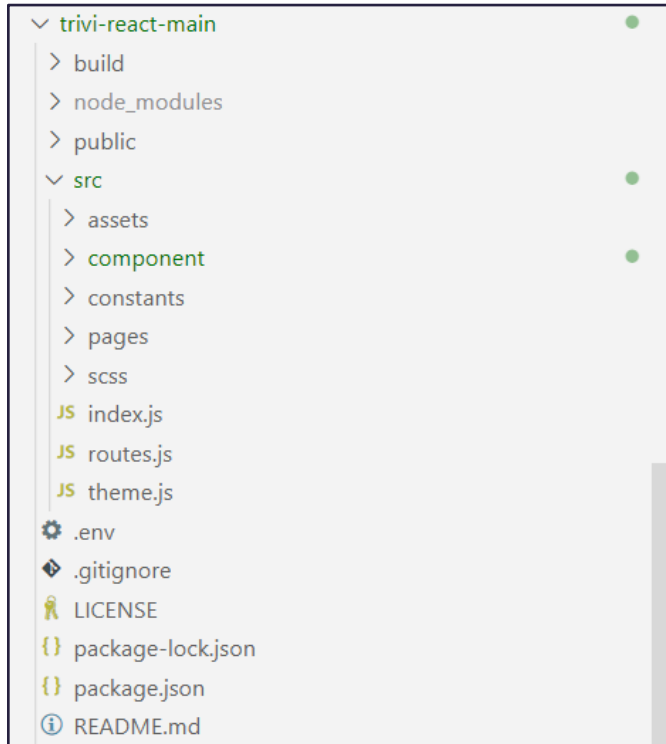
Hình 22. Cấu trúc quản lý phía hệ thống gợi ý sản phẩm

| Folder/File | Mô tả |
|----------------------------|---|
| Airflow/dags/scheduler_job | Chứa toàn bộ phần nhiệm vụ của máy chủ lập lịch |

| Folder/File | Mô tả |
|---|---|
| Intergrate_script/script.js | Chứa đoạn mã các trang TMĐT cần nhúng vào. |
| recommender-backend-main | Tập gốc, chứa toàn bộ phần xây dựng bộ phụ trợ, được phát triển trên Django |
| recommender-backend-main/django | Tập chứa các phần quản lý môi trường ảo dành cho máy chủ gọi ý. |
| manage.py | Bộ lưu trữ các chức năng thực thi (runserver, migrate) |
| recommender/static | Chứa các file tĩnh: CSS, js sẽ hiển thị cho người dùng |
| recommender/recommender | Thư mục ứng dụng mặc định, lưu trữ tất cả file cấu hình của dự án |
| recommender/model | Chứa tất cả các mô hình đã được huấn luyện |
| recommender/dimadb | App Django liên quan đến chức năng gợi ý sản phẩm, bắt sự kiện, cung cấp dữ liệu vẽ biểu đồ, v.v. |
| recommender/core | App Django liên quan đến chức năng đến đăng ký, đăng nhập, authentication |
| ./dimadb/personalize_recommendation.py | Tập tin chứa mã tất cả các chức năng của thuật toán gợi ý. |
| _app/view.py | Chứa các hàm để nhận yêu cầu và phản hồi kết quả. |
| _app/models.py | Định nghĩa các bảng trong cơ sở dữ liệu |
| _app/urls.py | Ánh xạ giữa đường dẫn URL tới các hàm được cài đặt trong View.py |

Bảng 5. Mô tả các thư mục chính phía hệ thống gợi ý sản phẩm

4.3.2. Thành phần giao diện khách hàng doanh nghiệp



Hình 23. Cấu trúc quản lý phía giao diện khách hàng doanh nghiệp

| Folder/File | Mô tả |
|--|---|
| recommender-react-main | Tập gốc, chứa toàn bộ phần xây dựng phía giao diện khách hàng doanh nghiệp, được phát triển trên ReactJS. |
| recommender-react-main/package.json | Chứa các tên, phiên bản các packages của Javascript được sử dụng ở Front-end |
| trivi-react-main/node_modules | Chứa các packages của Javascript đã được cài đặt |
| trivi-react-main/src | Mã thực thi giao diện sẽ được cài đặt tại đây |
| trivi-react-main/src/assets | Chứa các hình ảnh sử dụng trên giao diện |

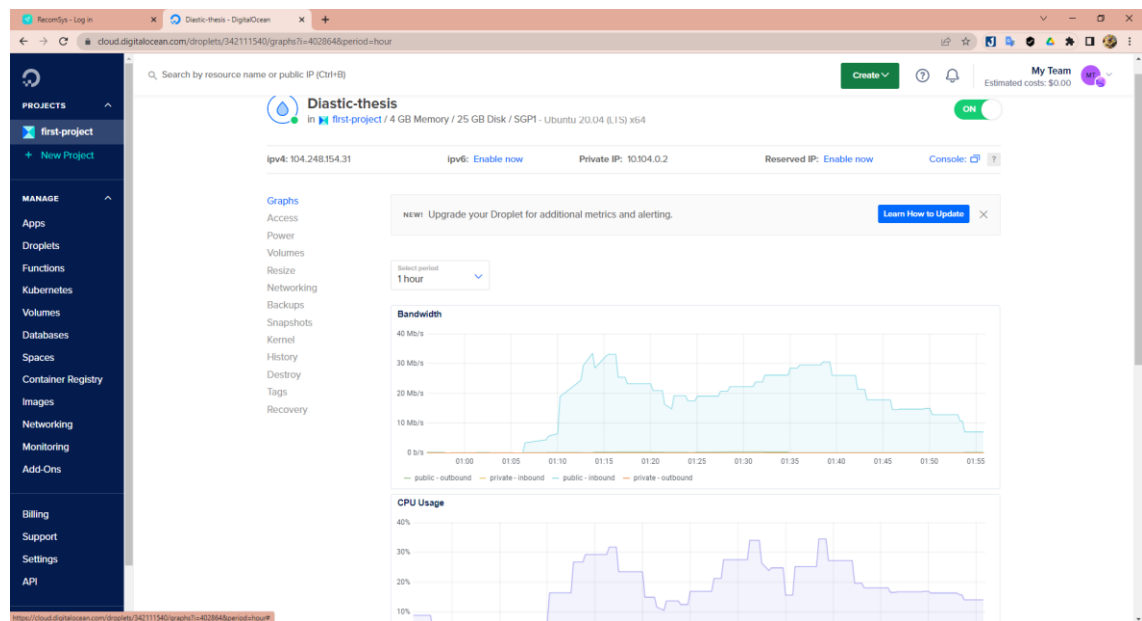
| Folder/File | Mô tả |
|---------------------------------------|---|
| trivi-react-main/src/component | Chứa mã cài đặt của các thành phần, chẳng hạn như chart, card, header, v.v. |
| trivi-react-main/src/contants | Chứa các hàm được sử dụng thường xuyên, các hằng số được sử dụng như địa chỉ IP, v.v. |
| trivi-react-main/src/pages | Chứa các file cài đặt giao diện cho từng trang tương ứng |
| trivi-react-main/src/scss | Chứa các file “css” định dạng giao diện |
| trivi-react-main/src/routes.js | Khai báo các router, có nhiệm vụ chuyển hướng URL trên thanh trình duyệt đến trang tương ứng |
| trivi-react-main/src/theme.js | Cài đặt chế độ hiển thị cho website, có 2 loại: dark theme (chế độ tối) và light theme (chế độ bình thường) |

Bảng 6. Mô tả thư mục chính phía giao diện khách hàng doanh nghiệp

4.3.3. Triển khai thành phần trên các máy chủ đám mây

- Digital Ocean - máy chủ cá nhân ảo (VPS)

Để có thể deploy một SaaS, chúng tôi cần có một máy chủ sau đó sao chép các đoạn code lên đó. Chúng tôi thuê một máy chủ Digital Ocean để thực hiện việc triển khai dịch vụ trong luận văn này.



Hình 24. Màn hình làm việc của Digital Ocean

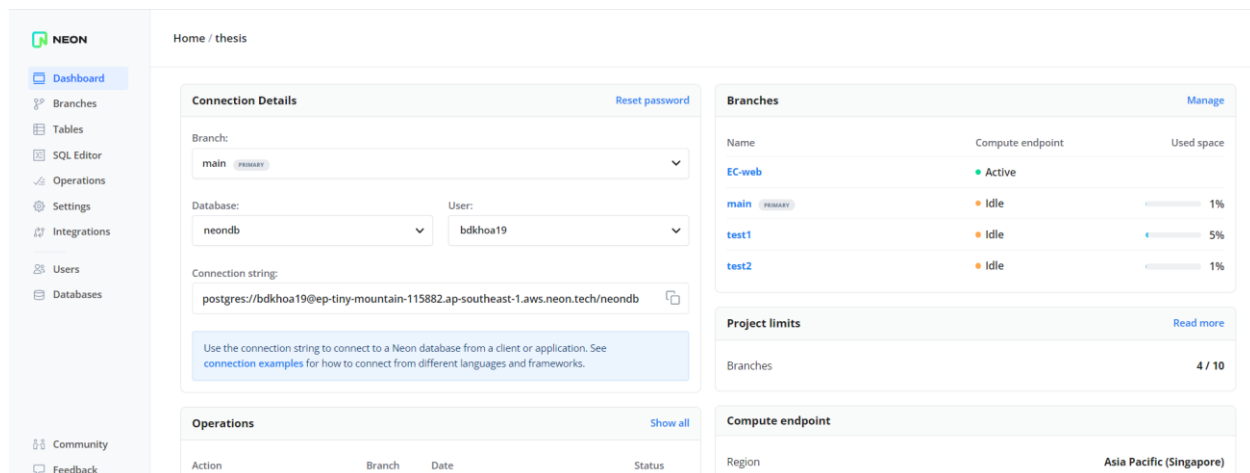
Để có thể sử dụng, chúng tôi kết nối với máy chủ theo IPv4 mà Digital Ocean đã cung cấp.

| | |
|---------------------|---------------|
| Hệ điều hành | Ubuntu 20.04 |
| Loại CPU | Regular Intel |
| vCPUs | 2vCPUs |
| Ram | 4GB |
| Ổ cứng SSD | 25GB |

Bảng 7. Thông tin server thuê từ Digital Ocean

- Neon - Serverless Postgres (Cơ sở dữ liệu không cần máy chủ)

Theo như kiến trúc luận văn ở phần 3.2.3, chúng tôi sử dụng Neon để cung cấp các cơ sở dữ liệu.



Hình 25. Màn hình làm việc của Neon

Chúng tôi chia làm 4 cơ sở dữ liệu chính, với mỗi cơ sở dữ liệu sẽ thuộc vào một nhánh.

| CSDL | Nhiệm vụ |
|---------------------|---|
| Main | Cơ sở dữ liệu quản lý cho Recommender server |
| test1, test2 | Lần lượt là các cơ sở dữ liệu sẽ cung cấp và thực hiện dịch vụ cho các trang TMĐT |
| EC-web | CSDL của trang TMĐT mà chúng tôi sẽ ví dụ. |

Bảng 8. Các CSDL của hệ thống

Để có thể truy cập vào cơ sở dữ liệu đã cung cấp, chúng tôi cần phải sử dụng đúng endpoint ở mỗi nhánh mà Neon cung cấp cùng với username và password.

4.4. Thực nghiệm

Phần này nhằm kiểm tra tính khả thi của luận văn. Chúng tôi tiến hành kiểm chứng trên 2 website TMĐT hoàn toàn khác nhau về cấu trúc và công nghệ:

- Xây dựng trên nền tảng NodeJS (back-end) và ReactJS (front-end)
- Dựa trên framework của Wordpress

Chúng tôi làm thực nghiệm này với mục đích chứng minh hệ thống dịch vụ tương tác của chúng tôi có thể tích hợp tốt trên đa dạng nền tảng công nghệ. Thông qua thực nghiệm, chúng tôi càng thấy rõ vấn đề khi can thiệp vào giao diện website TMĐT, chúng tôi không đảm bảo pop-up mà chúng tôi đưa ra phù hợp với thiết kế của doanh nghiệp, vô tình làm mất đi tính đặc trưng vốn có của doanh nghiệp.

4.3.1. Kiểm chứng trên website TMĐT

Sau khi đã hoàn tất quá trình thiết lập và cấu hình máy chủ, chúng tôi chạy máy chủ gợi ý và giao diện khách hàng doanh nghiệp theo kiến trúc được đề ra. Tiếp theo, để kiểm chứng mô hình gợi ý sản phẩm hướng dịch vụ, chúng tôi đã sử dụng một website thương mại điện tử mẫu và nhúng đoạn script vào source code của website.

```
<div id="recommendations"></div>
<link rel="stylesheet" href="http://localhost:8000/static/dimadb/css/recommender.css">
<script src="http://localhost:8000/static/dimadb/js/recommender.js"></script>
<script>
  var evnts = ["click", "focus", "blur"];
  var currentUrl = window.location.href;
  var tmp = "";

  window.addEventListener('load', function () {showPopup_onscreen()})

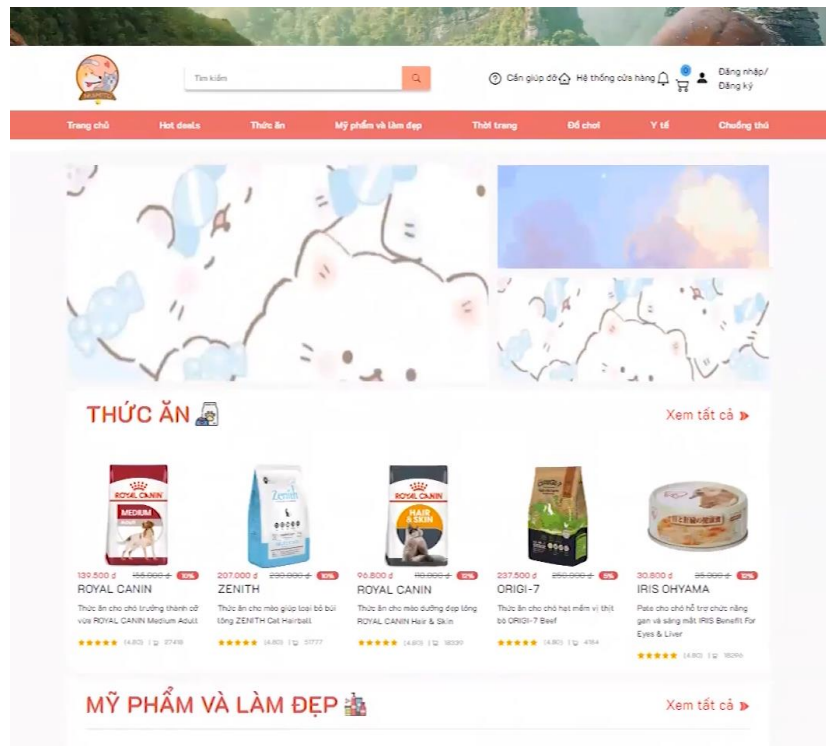
  const func_debounce = debounce_leading(function(e) {
    tmp = currentUrl
    if (currentUrl != window.location.href) currentUrl = window.location.href

    var anchor = getParentAnchor(e.target),nextUrl = "";
    if(anchor !== null) nextUrl = anchor.href;

    capture_event(e,tmp,nextUrl);
  }, 200);

  for (var i = 0; i < evnts.length; i++) {
    document.addEventListener("" + evnts[i] + "", func_debounce);
  }
</script>
```

Hình 26. Đoạn mã được nhúng vào website TMĐT (sử dụng ReactJS)



Hình 27. Trang chủ website TMĐT mẫu (ReactJS)

Hệ thống của chúng tôi cung cấp đa dạng thuật toán gợi ý sản phẩm đã trình bày ở phần 2.3, khi doanh nghiệp đã cung cấp đầy đủ dữ liệu về sản phẩm và người dùng, việc áp dụng thuật toán có mức độ ưu tiên như sau:

- Mức độ 1: Đối với người dùng chưa đăng nhập website TMĐT, chúng tôi sẽ xem họ như người dùng mới và áp dụng thuật toán Content-based hoặc Hot items để gợi ý sản phẩm.
- Mức độ 2: Khi người đã đăng nhập vào website TMĐT, chúng tôi áp dụng thuật toán đã cài đặt theo nhu cầu của khách hàng doanh nghiệp trên từng trang của website TMĐT. Việc cài đặt này thực hiện ở giao diện quản lý mà chúng tôi sẽ trình bày ở phần 4.3.2.

Sau đây, chúng tôi đóng vai là khách hàng doanh nghiệp của website TMĐT và đăng nhập vào website. Sau đó, chúng tôi lần lượt thực hiện các thao tác click, scroll, thêm vào giỏ hàng và chờ đợi kết quả gợi ý sản phẩm. Sau khi bắt được

các sự kiện trên website TMĐT, hệ thống gợi ý sản phẩm sẽ phản hồi và lưu dưới localStorage danh sách sản phẩm gợi ý với cấu trúc JSON được minh họa qua hình 28:

```
▼ {type: 'colab', message: 'popup', list_recommend: Array(3)} 1
  ▼ list_recommend: Array(3)
    ► 0: {id: '200081', name: 'Khay vệ sinh cho mèo MAKAR Cat Litter Station', url: 'http://104.248.154.31:3000/produc
    ► 1: {id: '200031', name: 'Nhà cây cho mèo Cat Tree QQ80862', url: 'http://104.248.154.31:3000/product/200031', im
    ► 2: {id: '200068', name: 'Kềm cắt bấm móng chó mèo DELE Q022 Pet Nail Clipper', url: 'http://104.248.154.31:3000/
      length: 3
    ► [[Prototype]]: Array(0)
      message: "popup"
      type: "colab"
    ► [[Prototype]]: Object
```

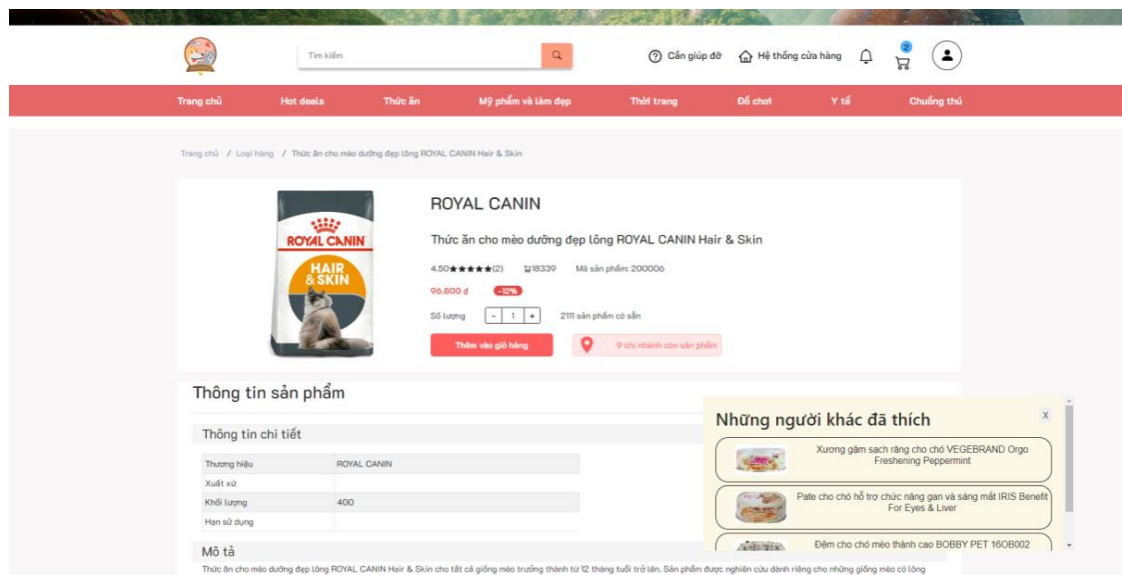
Hình 28. Cấu trúc JSON kết quả gợi ý sản phẩm

Cấu trúc JSON được mô tả qua bảng 9 như sau:

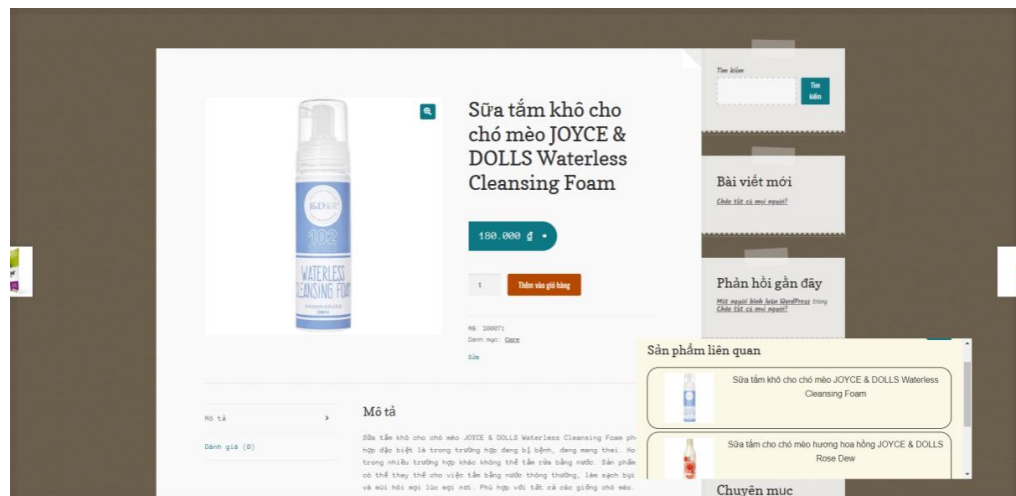
| Field | Mô tả |
|-----------------------|--|
| type | Loại thuật toán đang sử dụng |
| list_recommend | Danh sách kết quả, bao gồm id – mã sản phẩm, name – tên sản phẩm, url – đường dẫn đến sản phẩm, image – hình ảnh |

Bảng 9. Mô tả cấu trúc JSON kết quả gợi ý sản phẩm

Để trực quan hơn cho việc hiển thị kết quả, chúng tôi minh họa theo phương thức 2 đã nhắc đến ở phần 3.3.1, kết quả sản phẩm gợi ý sẽ được hiển thị như hình 29:



Hình 30. Kết quả minh họa trên website TMĐT (ReactJS)



Hình 29. Kết quả minh họa trên website TMĐT (Wordpress)

Tại góc phía dưới bên phải màn hình, xuất hiện một cửa sổ pop-up hiển thị thông tin về các sản phẩm được gợi ý theo thuật toán. Cửa sổ pop-up này sẽ tự động biến mất sau 20 giây, hoặc người dùng có thể đóng nó bằng cách nhấn vào nút tắt ở góc phía trên bên phải của cửa sổ pop-up. Chúng tôi áp dụng các thuật

toán gợi ý đã trình bày ở phần 2.3, trên mỗi thuật toán khác nhau, pop-up sẽ hiển thị tiêu đề tương ứng như hình 30, cụ thể là:

- Thuật toán Collaborative Filtering: “Có thể bạn sẽ thích”
- Thuật toán Demographic: “Những người khác đã thích”
- Thuật toán Content-based: “Sản phẩm liên quan”
- Thuật toán Hot items: “Sản phẩm nổi bật”




Hình 31. Minh họa pop-up hiển thị kết quả gợi ý sản phẩm

Với mỗi lần tương tác, người dùng sẽ nhận được kết quả gợi ý. Đồng thời, chúng tôi cũng lưu trữ các lần tương tác của người dùng và hiển thị ở phần Captured Data của phân hệ quản lý để giúp khách hàng doanh nghiệp có cái nhìn toàn vẹn về tương tác của người dùng.

4.3.2. Minh họa phân hệ quản lý

Dữ liệu tương tác của người dùng được chúng tôi lưu trữ và hiển thị ở giao diện phân hệ quản lý. Ngoài ra, trang quản lý của chúng tôi còn cho phép khách hàng doanh nghiệp cài đặt thông số cho hệ thống gợi ý sản phẩm. Khách hàng doanh nghiệp cần đăng ký tài khoản để sử dụng hệ thống của chúng tôi thông qua giao diện như hình 32:

Log in



Log in

Create new account? [Sign up](#)

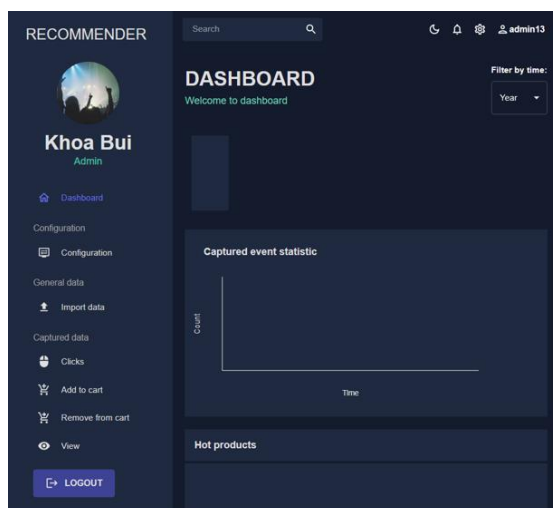
Create an account

☐ I agree to the terms and conditions

Sign up

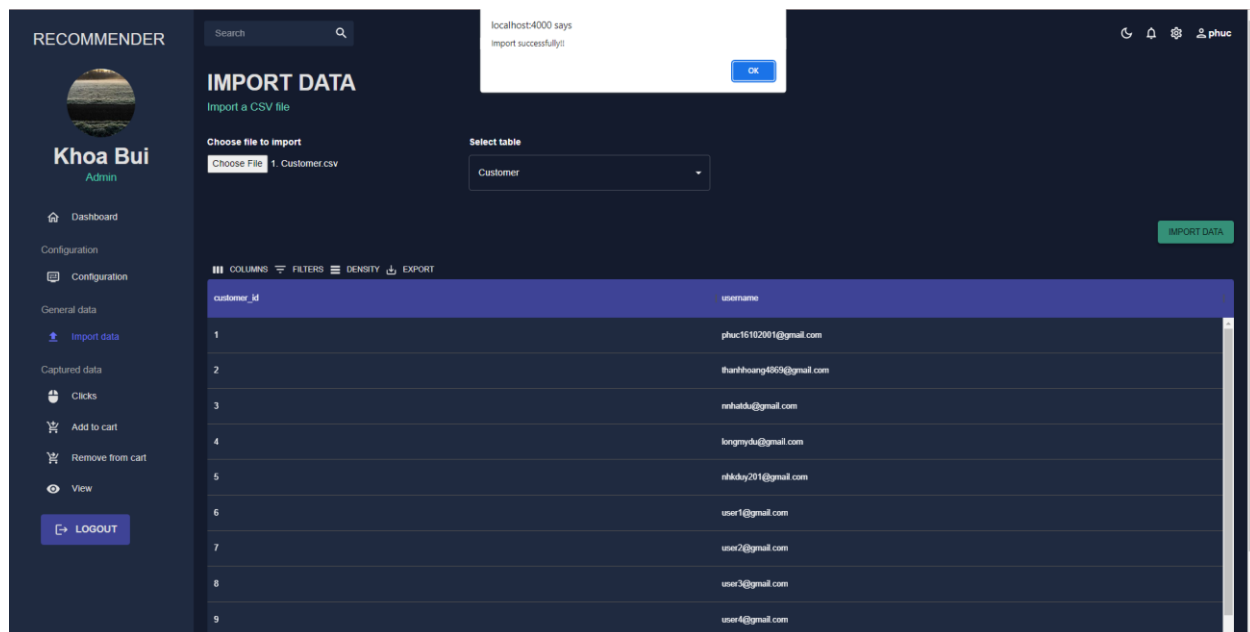
Hình 32. Giao diện đăng nhập (bên trái), đăng ký tài khoản (bên phải)

Do chúng tôi chỉ cung cấp hạn chế số lượng CSDL, vì vậy, khi hết CSDL có sẵn, chúng tôi sẽ thông báo không đủ CSDL cung cấp và không cho doanh nghiệp đăng ký tài khoản mới. Khi đăng nhập thành công, doanh nghiệp sẽ được chuyển hướng đến trang quản lý hệ thống gợi ý sản phẩm:



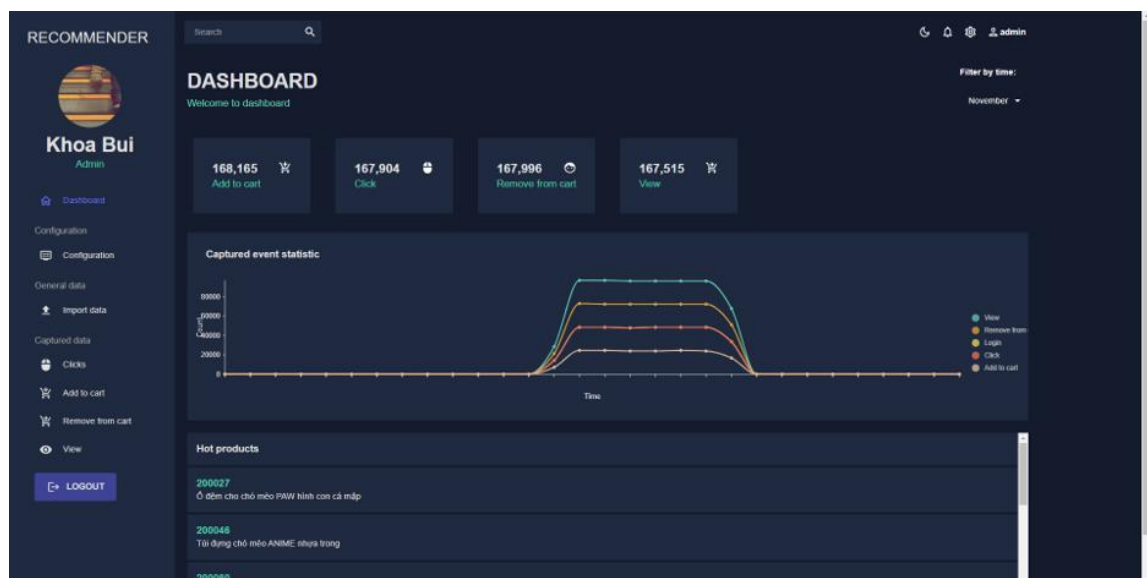
Hình 33. Giao diện Dashboard (khi chưa có dữ liệu)

Đối với những khách hàng doanh nghiệp vừa mới tạo tài khoản, cần nhập các thông tin về người dùng và sản phẩm thông qua Tab “Import Data” để hệ thống có dữ liệu cho việc huấn luyện mô hình gợi ý.



Hình 34. Giao diện Import data.

Tính đến lúc này, hệ thống đã có thể bắt các sự kiện diễn ra trên website TMĐT. Khi đó, ở giao diện Dashboard sẽ hiển thị thông tin thống kê như hình 25:

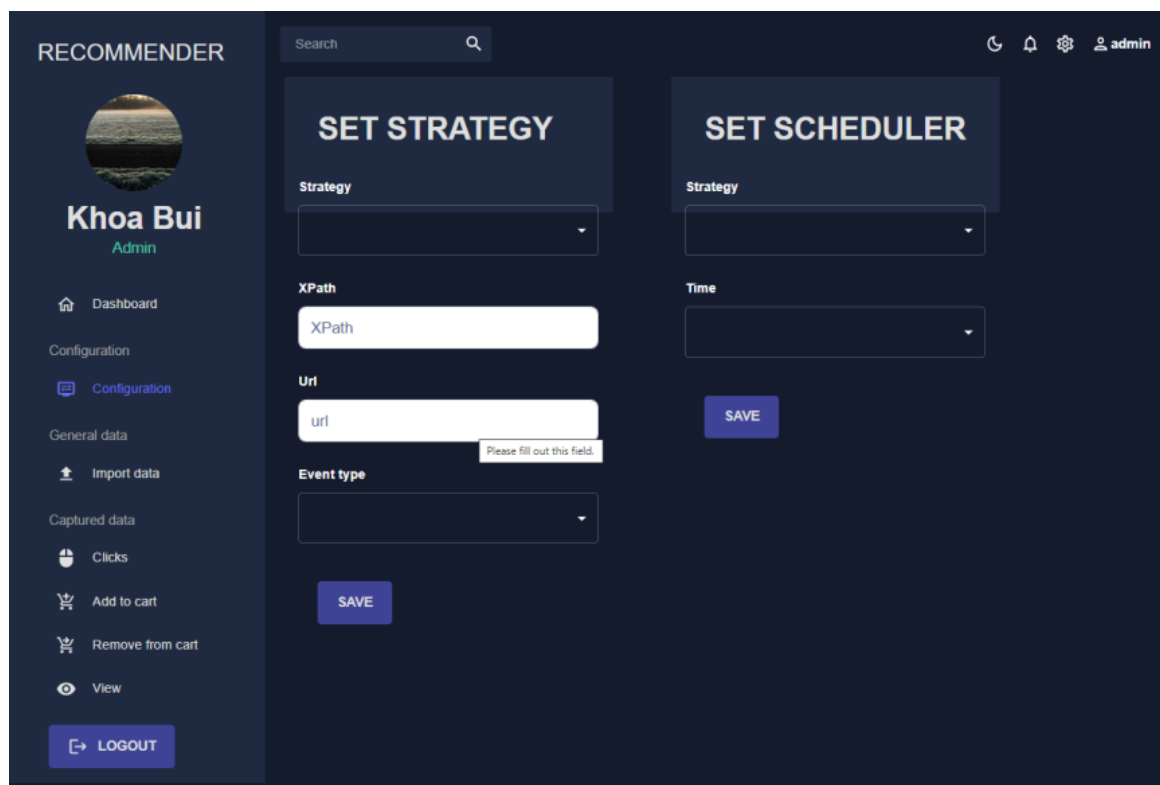


Hình 35. Giao diện Dashboard

Hình 35 chứa các thành phần chính như sau:

- Tổng số hành vi bắt được theo từng loại sự kiện
- Biểu đồ đường thống kê sự kiện theo tháng
- Top 10 sản phẩm phổ biến

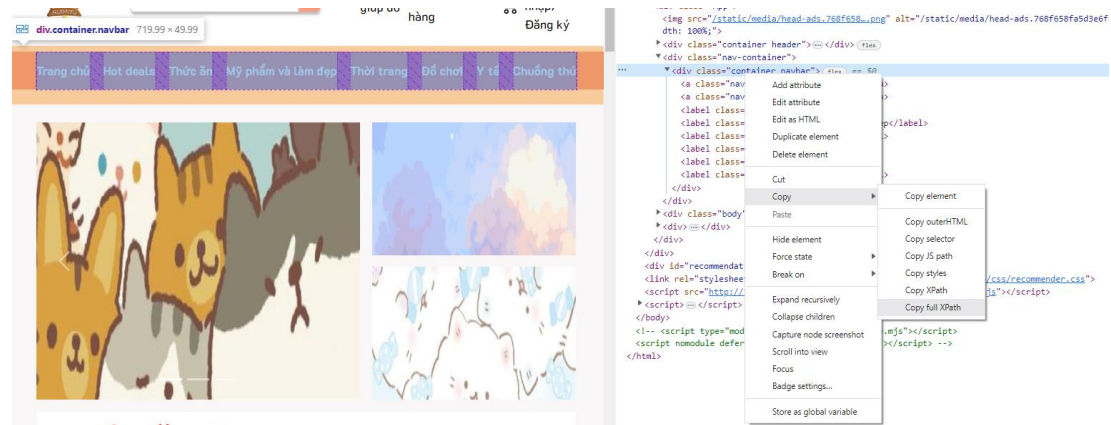
Như chúng tôi đã trình bày cách thức hoạt động ở các phần trên, doanh nghiệp có thể tùy chỉnh chiến lược (thuật toán) cho từng trang và cài đặt chu kỳ huấn luyện mô hình thông qua Tab “Configuration”:



Hình 36. Giao diện Configuration

Đối với cài đặt chiến lược (thuật toán) gợi ý sản phẩm, doanh nghiệp cần cung cấp XPath đến phần tử cần bắt sự kiện. Cách lấy XPath như sau: Tại màn hình cần bắt sự kiện, doanh nghiệp nhấn F12 hoặc bấm chuột phải chọn Inspect để mở chế độ Developer view. Tại đây, doanh nghiệp trở đến thành phần cần bắt

sự kiện, bấm chuột phải và chọn Copy full XPath. Chúng tôi minh họa trực quan qua hình 37:



Hình 37. Cách lấy XPath

Ngoài ra, dữ liệu về tương tác của người dùng trên hệ thống được chúng tôi lưu trữ và hiển thị dưới dạng bảng biểu ở mục Captured data. Doanh nghiệp có thể xem dữ liệu và filter theo trường có nhu cầu. Hình 38 minh họa giao diện hiển thị dữ liệu của hành vi Click:

The image shows a dashboard titled 'CLICK STATISTICS' with a table of click events. The table has columns for product_id, product_name, count, and price. The data is filtered by time, showing a list of products and their associated click counts and prices.

| product_id | product_name | count | price |
|------------|---|-------|---------|
| 200000 | Thức ăn cho chó Poodle con ROYAL CANIN Poodle Puppy | 2076 | 165,000 |
| 200001 | Thức ăn cho chó con cỡ nhỏ ROYAL CANIN Mini Puppy | 2001 | 105,000 |
| 200002 | Thức ăn cho mèo trưởng thành ROYAL CANIN Regular Fit 32 | 2036 | 95,000 |
| 200003 | Thức ăn cho mèo trưởng thành ROYAL CANIN Indoor 27 | 2027 | 105,000 |
| 200004 | Thức ăn cho chó trưởng thành cỡ vừa ROYAL CANIN Medium Adult | 2000 | 165,000 |
| 200005 | Thức ăn cho mèo giúp loại bỏ lông ZENTH Cat Hairball | 2040 | 230,000 |
| 200006 | Thức ăn cho mèo dưỡng da ROYAL CANIN Hair & Skin | 2030 | 110,000 |
| 200007 | Thức ăn cho chó hạt miễn dịch bả CRIGI 7 Beef | 2053 | 250,000 |
| 200008 | Pate cho chó hỗ trợ chức năng gan và sáng mắt BRS Benefits For Eyes & Liver | 2046 | 35,000 |
| 200009 | Pate cho chó vị bò và rau BRS Benefits Beef Vegetable | 1990 | 85,000 |
| 200010 | Pate cho chó giảm rụng lông BRS Benefits For Fur | 2021 | 35,000 |

Hình 38. Giao diện xem dữ liệu hành vi Click

4.5. Kết luận chương

Chương 4 đã trình bày giai đoạn chuẩn bị khởi tạo môi trường lập trình cũng như minh họa kết quả thông qua thực nghiệm của chúng tôi. Hệ thống của chúng tôi chỉ mới được kiểm tra trên các website TMĐT được xây dựng từ NodeJS (back-end) kết hợp ReactJS (front-end) và website dựa trên framework của Wordpress. Chúng tôi sẽ tiếp tục kiểm thử khả năng tích hợp hệ thống gợi ý sản phẩm vào các môi trường khác trong tương lai.

CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Chương 5 trình bày kết luận, đóng góp của luận văn, cũng như các đánh giá về quá trình hoàn thiện luận văn. Chúng tôi trước hết trình bày cụ thể kết quả đạt được của luận văn về mặt lý thuyết cũng như thực nghiệm. Các kiến thức đạt được cũng như những hạn chế chưa khắc phục được. Cuối cùng chúng tôi trình bày hướng phát triển của đề tài trong tương lai.

5.1. Kết quả đạt được

Quá trình nghiên cứu và thực hiện luận văn giúp chúng tôi đạt được những kết quả kết quả như sau:

- Tìm hiểu và sử dụng các công cụ đã trình bày ở phần 2.4 cho phát triển dự án.

Từ luận văn này, chúng tôi có thể phối hợp sử dụng các công cụ trên, ngoài ra, chúng tôi biết thêm về triển khai hệ thống trên đám mây thay vì sử dụng máy chủ truyền thống.

- Tìm hiểu về SaaS, dịch vụ tương tác.

Chúng tôi nắm được khái niệm về SaaS, dịch vụ tương tác cũng như mô hình, cách thức hoạt động của chúng. Luận văn này giúp chúng tôi nhận ra tầm quan trọng của SaaS và dịch vụ tương tác trong việc cung cấp giải pháp kinh doanh hiệu quả nhằm đáp ứng nhu cầu ngày càng cao của người dùng.

- Tìm hiểu về các thuật toán gợi ý sản phẩm.

Chúng tôi phân loại và nhận biết sự khác biệt giữa các thuật toán gợi ý sản phẩm. Các thuật toán gợi ý sản phẩm ngày càng được áp dụng rộng rãi trong thị trường TMĐT hiện nay.

- Xây dựng mô hình cơ sở dữ liệu và đề xuất kiến trúc hệ thống.

Chúng tôi cung cấp mô hình CSDL với các đối tượng cần thiết cho hệ thống gợi ý sản phẩm. Ngoài ra, chúng tôi đề xuất kiến trúc hệ thống nhằm xác định các thành phần của hệ thống và mối quan hệ, tương tác của chúng trong hệ thống.

- Tích hợp thành công dịch vụ gợi ý sản phẩm vào website TMĐT.

Chúng tôi đã tiến hành kiểm thử trên nhiều website TMĐT mẫu và nhận được sự phản hồi từ hệ thống gợi ý sản phẩm.

- Cung cấp giao diện quản lý

Chúng tôi cung cấp các chức năng như đăng ký, đăng nhập, xem báo cáo thống kê, nhập dữ liệu, cài đặt thông số cho hệ thống gợi ý sản phẩm.

5.2. Hạn chế của luận văn

Do hạn chế về thời gian, luận văn còn vài điểm cần khắc phục:

Về tính hiệu quả thuật toán: Chúng tôi chủ yếu tập trung vào việc tích hợp dịch vụ gợi ý sản phẩm mà chưa thể đánh giá về tính hiệu quả của thuật toán. Việc đánh giá tính hiệu quả còn đòi hỏi phải đánh giá ảnh hưởng của dịch vụ lên trải nghiệm của người dùng và kết quả kinh doanh của doanh nghiệp. Luận văn này chỉ là bản mẫu (prototype) của dịch vụ và chỉ mới được kiểm thử trên dữ liệu mẫu mà chúng tôi cung cấp. Do đó, hiệu quả của thuật toán gợi ý vẫn còn đang trong quá trình đánh giá và cải thiện.

Về kiến trúc: Chúng tôi chỉ có thể cung cấp đồng thời số lượng hạn chế CSDL cho doanh nghiệp. Khi hết CSDL có sẵn, chúng tôi ngừng tiếp nhận đăng ký sử dụng dịch vụ từ phía doanh nghiệp.

Về hiệu suất: Về phân hệ quản lý hệ thống, chúng tôi gặp vấn đề xử lý dữ liệu lớn, tải trang chậm. Điều này gây ra hiệu suất kém, có thể làm ảnh hưởng đến trải nghiệm của doanh nghiệp khi sử dụng, do đó, chúng tôi vẫn trong quá trình cải thiện hiệu suất.

Về tính khả dụng: Hệ thống của chúng tôi chỉ áp dụng trên những đối tượng mà doanh nghiệp cung cấp, nghĩa là, chúng tôi chỉ xét trên những đối tượng đã được khai báo XPath, URL và thuật toán áp dụng.

5.3. Hướng phát triển tương lai

Mô hình tương tác hướng dịch vụ cho hệ thống gợi ý sản phẩm là một phương pháp tiềm năng để cải thiện trải nghiệm người dùng và tăng tính hiệu quả của các trang TMĐT của các doanh nghiệp vừa và nhỏ. Điều đó cũng cho thấy tiềm năng của mô hình này về việc tạo ra các dịch vụ đột phá và làm thay đổi cách mà doanh nghiệp tương tác với người dùng. Chúng tôi cảm thấy việc tích hợp mô hình này vào website TMĐT của các doanh nghiệp vừa và nhỏ là một hướng đi mới và chưa phổ biến trên thị trường. Ở phần trình bày này, chúng tôi sẽ đề xuất một số định hướng phát triển để có thể hoàn toàn đáp ứng được nhu cầu của các doanh nghiệp vừa và nhỏ.

5.3.1. Phát triển kiến trúc

Trong luận văn này, vì hạn chế về mặt thiết kế kiến trúc đã nói ở phần 5.2, hệ thống của chúng tôi chỉ có thể cung cấp được cho một số ít doanh nghiệp đồng thời. Trong lĩnh vực TMĐT, các ngày "super sale" có thể gây ra tình trạng quá tải hệ thống. Để khắc phục vấn đề này, chúng tôi đang xem xét các hướng phát triển mới cho kiến trúc hệ thống.

Về mặt kiến trúc hệ thống: Chúng tôi đang xem xét thiết kế theo hướng microservice để có thể cung cấp dịch vụ cho nhiều đơn vị doanh nghiệp hơn. Điều này sẽ giúp giảm bớt tình trạng quá tải khi có nhiều yêu cầu đồng thời.

Về mặt cơ sở dữ liệu: Hiện tại, chúng tôi lưu tất cả trên cơ sở dữ liệu quan hệ. Tuy nhiên, với sự phát triển của NoSQL, chúng tôi có thể tối ưu hóa một số phần bằng cơ sở dữ liệu NoSQL để giúp giảm tải cho hệ thống. Một ví dụ điển hình là việc lưu trữ session bằng Redis (Key-value database). Với số lượng đơn hàng lớn trong các ngày mua sắm, cơ sở dữ liệu key-value có thể giải quyết vấn

để mở rộng cho lượng dữ liệu khổng lồ thông qua quá trình xử lý và lưu trữ phân tán.

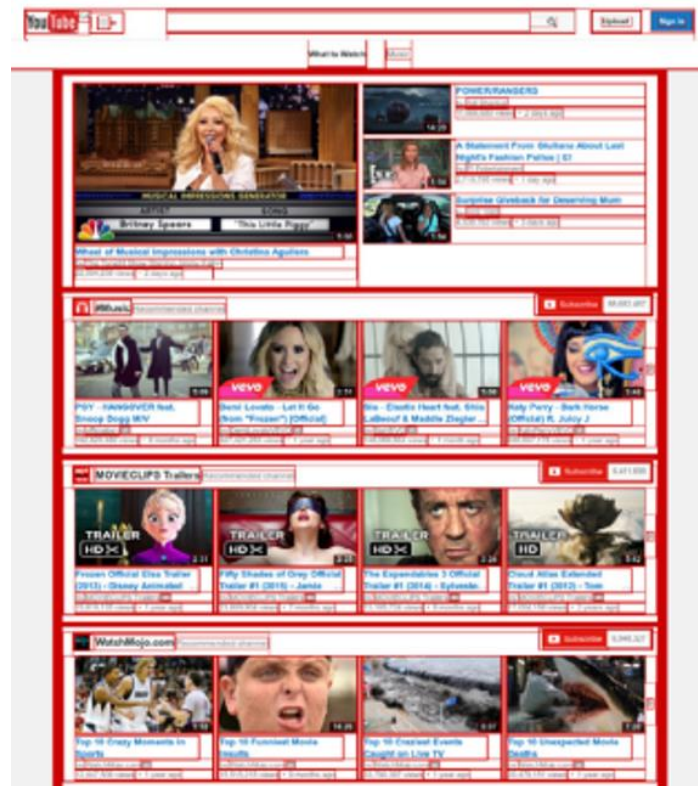
5.3.2. Phát triển về hiệu quả thuật toán

Vì tính hiệu quả thuật toán chưa được đề cập nhiều, chúng tôi sẽ tiếp tục đánh giá và cải thiện tính hiệu quả thuật toán trên các website TMĐT thực tế. Để phát triển thuật toán gợi ý sản phẩm một cách nâng cao, chúng tôi đã đề xuất vài hướng đi tiếp theo như mở rộng thuật toán để bao gồm Hành trình khách hàng (customer journey) hoặc sử dụng phương pháp hybrid.

Về mở rộng và tận dụng, hệ thống mô hình tương tác hướng dịch vụ có thể trở thành một công cụ đắc lực giúp các kỹ sư máy học thử nghiệm và phát triển các thuật toán mới trên các website TMĐT thực tế mà không gây ảnh hưởng đến trang website đó. Điều này sẽ giúp nâng cao chất lượng và hiệu quả của quá trình gợi ý sản phẩm trên thị trường TMĐT.

5.3.3. Phát triển về cách bắt sự kiện

Hiện nay, rất nhiều luận văn đã tập trung vào việc nghiên cứu và hiểu cấu trúc của các website thông qua phương pháp phân đoạn trang web (Web Page Segmentation). Chúng tôi đang xem xét sử dụng thuật toán này để hỗ trợ khách hàng doanh nghiệp trong việc chọn các phần trên website để bắt sự kiện, đồng thời thuật toán này có thể cung cấp cho chúng tôi các đường dẫn XPath một cách dễ dàng và chính xác.



Hình 39. Minh họa mô hình Web Page Segmentation [21]

5.3.4. Kết luận

Việc triển khai mô hình gợi ý sản phẩm đòi hỏi sự tích hợp chặt chẽ giữa các dịch vụ khác nhau, điều này có thể không dễ dàng cho các doanh nghiệp vừa và nhỏ, đặc biệt là trong bối cảnh nguồn lực có hạn. Với các định hướng phát triển mà chúng tôi nêu ở trên, chúng tôi tin rằng mô hình tương tác hướng dịch vụ cho hệ thống gợi ý sản phẩm có thể mang lại giá trị lớn cho các doanh nghiệp vừa và nhỏ ở một tương lai không xa.

PHỤ LỤC

Mô tả thuộc tính của cơ sở dữ liệu

Mô tả thuật ngữ trong tài liệu rất quan trọng để đảm bảo tính tái sử dụng và hiệu quả truyền tải thông tin. Phần dưới đây trình bày chi tiết về thuộc tính và quan hệ của các đối tượng trong mô hình được đề xuất:

| CUSTOMER | |
|--------------------|---|
| Thuộc tính | Mô tả |
| customer_id | ID, định danh người dùng của website thương mại điện tử |
| ip_address | Địa chỉ ip của người dùng |
| username | Tên đăng nhập vào website thương mại điện tử của người dùng |

Bảng 10. Mô tả bảng CUSTOMER

| CUSTOMER_PROFILE | |
|--------------------|---|
| Thuộc tính | Mô tả |
| customer_id | ID, định danh người dùng của website thương mại điện tử |
| dob | Ngày tháng năm sinh của người dùng |
| gender | Giới tính người dùng |
| city | Thành phố nơi người dùng sinh sống |
| country | Quốc gia nơi người dùng sinh sống |
| first_name | Tên của người dùng |

| CUSTOMER_PROFILE | |
|-------------------------|------------------------------|
| Thuộc tính | Mô tả |
| last_name | Họ của người dùng |
| email | Địa chỉ email của người dùng |
| phone | Số điện thoại của người dùng |

Bảng 11. Mô tả bảng CUSTOMER_PROFILE

| PRODUCT | |
|---------------------|---|
| Thuộc tính | Mô tả |
| product_id | ID, định danh sản phẩm của website thương mại điện tử |
| product_name | Tên sản phẩm |
| price | Giá niêm yết của sản phẩm |
| revenue | Giá bán của sản phẩm |
| category | Phân loại sản phẩm |
| discription | Mô tả sản phẩm |
| url | Link sản phẩm của website thương mại điện tử |

Bảng 12. Mô tả bảng PRODUCT

| SESSION | |
|-------------------|---|
| Thuộc tính | Mô tả |
| Session_id | ID, định danh phiên đăng nhập của người dùng trong trang website thương mại điện tử |

| SESSION | |
|-------------|---|
| Thuộc tính | Mô tả |
| Start_time | Thời điểm bắt đầu phiên |
| End_time | Thời điểm kết thúc phiên |
| Customer_id | ID của người dùng, tham chiếu đến customer_id của bảng CUSTOMER |

Bảng 13. Mô tả bảng SESSION

| WEB_EVENT | |
|------------|---|
| Thuộc tính | Mô tả |
| Event_id | ID, định danh sự kiện bắt được từ web thương mại điện tử |
| Created_at | Thời điểm bắt sự kiện |
| Event_type | Loại sự kiện: click, view, add to cart, remove from cart |
| Session_id | ID của phiên đăng nhập, tham chiếu đến session_id của bảng CUSTOMER_SESSION |

Bảng 14. Mô tả bảng WEB_EVENT

| BUSINESS_TRANSACTION | |
|----------------------|--|
| Thuộc tính | Mô tả |
| transation_id | ID, định danh sự kiện bắt được từ web thương mại điện tử |
| Created_at | Thời điểm bắt sự kiện |

| BUSINESS_TRANSACTION | |
|-----------------------------|---|
| Thuộc tính | Mô tả |
| Transaction_revenue | Loại sự kiện: click (view), add to cart, remove from cart |
| Transation_shipping | ID của phiên đăng nhập, tham chiếu đến session_id của bảng CUSTOMER_SESSION |
| Transaction_status | Trạng thái của giao tác, bao gồm: Success, Cancel, Delivering |
| Session_id | ID của phiên đăng nhập, tham chiếu đến session_id của bảng CUSTOMER_SESSION |

Bảng 15. Mô tả bảng BUSINESS_TRANSACTION

| EVENT_ITEM | |
|-------------------|---|
| Thuộc tính | Mô tả |
| Event_id | ID sự kiện, tham chiếu đến event_id của bảng WEB_EVENT |
| Product_id | ID sản phẩm, tham chiếu đến product_id của bảng PRODUCT |
| Quantity | Số lượng sản phẩm |
| Price | Giá của sản phẩm |

Bảng 16. Mô tả bảng EVENT_ITEM

| TRANSACTION_ITEM | |
|-------------------------|--|
| Thuộc tính | Mô tả |
| transaction_id | ID giao tác, tham chiếu đến transaction_id của bảng BUSINESS_TRANSACTION |
| Product_id | ID sản phẩm, tham chiếu đến product_id của bảng PRODUCT |
| Quantity | Số lượng sản phẩm |
| Price | Giá niêm yết của sản phẩm |
| revenue | Giá bán của sản phẩm |

Bảng 17. Mô tả bảng TRANSACTION_ITEM

| SCHEDULER | |
|---------------------|--|
| Thuộc tính | Mô tả |
| scheduler_id | ID của bộ lập lịch |
| cycle_time | Thời gian chạy thuật toán gợi ý, trong đó có 3 giá trị chính: 1 – 1 ngày, 7 – 1 tuần, 30 – 1 tháng |

Bảng 18. Mô tả bảng SCHEDULER

| RECOMMENDER_MODEL | |
|--------------------------|---------------------------------|
| Thuộc tính | Mô tả |
| model_id | ID mô hình |
| created_at | Thời gian tạo mô hình |
| model_type | Loại thuật toán mà mô hình chạy |
| model_path | Đường dẫn lưu kết quả chạy |

Bảng 19. Mô tả bảng RECOMMENDER_MODEL

| RECOMMENDER_STRATEGY | |
|-----------------------------|--|
| Thuộc tính | Mô tả |
| id | ID chiến lược |
| strategy | Tên chiến lược, bao gồm Collaborative Filtering, Content-based, Demographic, Hot items |
| event_type | Loại sự kiện áp dụng chiến lược, bao gồm: Click, View, Add to cart, Remove from cart |
| url | Đường dẫn trang áp dụng chiến lược |
| xpath | Xpath của thành phần áp dụng chiến lược |

Bảng 20. Mô tả bảng RECOMMENDER_STRATEGY

TÀI LIỆU THAM KHẢO

- [1] Thapliyal, Bijaya Laxmi and Rani, Neha and Bhadula, Rakesh Chandra and Devrani, Vinay and Mani, Ambica Prakash, Effects of e-Commerce in the Current Scenario of Marketing (October 14, 2020). International Journal of Management, 11 (9), 2020, pp. 770-777
- [2] Radhakrishnan, Vaishak. (2021). E-Commerce industry and its effect on the world today. International Research Journal on Advanced Science Hub. 3. 23-29. 10.47392/irjash.2021.026.
- [3] Sách trắng Thương mại điện tử Việt Nam năm 2022, Cục Thương mại điện tử và Kinh tế số (Bộ Công Thương)
- [4] Quy mô thị trường e-commerce Việt Nam đạt 39 tỉ USD vào năm 2025, tác giả Linh Chi, trích từ báo Forbes Việt Nam. Ngày truy cập: 19/12/2022. Đường dẫn: <https://forbes.vn/quy-mo-thi-truong-e-commerce-viet-nam-dat-39-ti-usd-vao-nam-2025>
- [5] K. Shah, A. Salunke, S. Dongare and K. Antala, "Recommender systems: An overview of different approaches to recommendations," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 2017, pp. 1-4, doi: 10.1109/ICIIECS.2017.8276172.
- [6] Le Dinh, Thang; Vu, Thi My Hang; Dam, Nguyen Anh Khoa; and Nguyen, Chan Nam, "Trivi: A Conceptual Framework for Customer Intelligence Systems for Small and Medium-sized Enterprises" (2022). PACIS 2022 Proceedings. 177
- [7] Peng-yu Lu, Xiao-xiao Wu and De-ning Teng, "Hybrid Recommendation

- Algorithm for E-Commerce Website," 2015 8th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 2015, pp. 197-200, doi: 10.1109/ISCID.2015.140.
- [8] T. Chen, Y. Liang, T. Huang, J. Huang and C. Liu, "Agricultural Product Recommendation Model and E-Commerce System based on CFR Algorithm," 2022 IEEE 2nd International Conference on Electronic Technology, Communication and Information (ICETCI), Changchun, China, 2022, pp. 931-934, doi: 10.1109/ICETCI55101.2022.9832175.
- [9] Y. Fan, Y. Shen and J. Mai, "Study of the Model of E-commerce Personalized Recommendation System Based on Data Mining," 2008 International Symposium on Electronic Commerce and Security, Guangzhou, China, 2008, pp. 647-651, doi: 10.1109/ISECS.2008.106.
- [10] S. Aleem, F. Ahmed, R. Batool and A. Khattak, "Empirical Investigation of Key Factors for SaaS Architecture," in IEEE Transactions on Cloud Computing, vol. 9, no. 3, pp. 1037-1049, 1 July-Sept. 2021, doi: 10.1109/TCC.2019.2906299.
- [11] G. Liu, "Research on independent SaaS platform," 2010 2nd IEEE International Conference on Information Management and Engineering, Chengdu, China, 2010, pp. 110-113, doi: 10.1109/ICIME.2010.5477498.
- [12] Tìm hiểu về SaaS, Salesforce. Ngày truy cập: 20/12/2022. Đường dẫn: <https://www.salesforce.com/in/saas/>
- [13] Thống kê về SaaS, Asiasoft. Ngày truy cập: 20/12/2022. Đường dẫn: <https://asiasoft.com.vn/2022/03/17/saas-la-gi-viet-nam-da-phat-trien-mo-hinh-saas-nhu-the-nao/>
- [14] Das, Debashis & Sahoo, Laxman & Datta, Sujoy. (2017). A Survey on

Recommendation System. International Journal of Computer Applications. 160. 6-10. 10.5120/ijca2017913081.

[15] Tìm hiểu collaborative filtering, MachineLearningCoBan. Ngày truy cập: 25/12/2022. Đường dẫn:

<https://machinelearningcoban.com/2017/05/24/collaborativefiltering/>

[16] Tìm hiểu thuật toán content-based, MachineLearningCoBan. Ngày truy cập: 25/12/2022. Đường dẫn:

<https://machinelearningcoban.com/2017/05/17/contentbasedrecommendersys/>

[17] Tổng hợp các khảo sát về CRM. Ngày truy cập: 10/01/2023. Đường dẫn:

<https://www.linkedin.com/pulse/20-crm-facts-you-should-know-2021-develop-your-private-limited/>

[18] Báo cáo thống kê tại Hội nghị thượng đỉnh về trải nghiệm khách hàng của Gartner 2018. Ngày truy cập: 10/01/2023. Đường dẫn:

<https://www.gartner.com/en/newsroom/press-releases/2018-02-19-gartner-says-25-percent-of-customer-service-operations-will-use-virtual-customer-assistants-by-2020>

[19] Báo cáo về ảnh hưởng của Chatbot trong lĩnh vực ngân hàng. Ngày truy cập: 10/01/2023. Đường dẫn: <https://www.juniperresearch.com/press/bank-cost-savings-via-chatbots-reach-7-3bn-2023>

[20] Amazon Recommendation System, Recostream. Ngày truy cập: 10/01/2023.

Đường dẫn: <https://recostream.com/blog/amazon-recommendation-system>

[21] Michael Cormier, Karyn Moffatt, Robin Cohen, Richard Mann, Purely vision-based segmentation of web pages for assistive technology, Computer Vision and Image Understanding, Volume 148, 2016, Pages 46-66, ISSN 1077-3142.