**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP HÀ NỘI**

**KHOA CÔNG NGHỆ THÔNG TIN**

**======\*\*\*======**

****

BÁO CÁO NGHIÊN CỨU KHOA HỌC

**Đề tài :**

**Xây dựng ứng dụng web sử dụng luật kết hợp để trợ giúp mua hàng**

|  |  |
| --- | --- |
| **GVHD:** | **Ths. Nguyễn Thanh Hải** |
| **Đề tài số:** | **42** |
| **Thành viên:** | **Cao Văn Sơn – chủ nhiệm đề tài** |
|  | **Phạm Hoàng Tiến** |

Hà nội, Năm 2022

BỘ CÔNG THƯƠNG

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP HÀ NỘI**

**BÁO CÁO TỔNG KẾT**

**ĐỀ TÀI NGHIÊN CỨU KHOA HỌC CỦA SINH VIÊN**

XÂY DỰNG ỨNG DỤNG WEB SỬ DỤNG LUẬT KẾT HỢP ĐỂ TRỢ GIÚP MUA HÀNG

Sinh viên thực hiện:

1. Cao Văn Sơn – 2020600313
2. Phạm Hoàng Tiến – 2020600235

Khoa: Công nghệ thông tin

Người hướng dẫn: Ths. Nguyễn Thanh Hải

LỜI MỞ ĐẦU

Trong những năm gần đây, sự phát triển của thương mại điện tử (E-Commerce) đã đem lại nhiều lợi ích to lớn cho nền kinh tế toàn cầu. Thông qua thương mại điện tử, nhiều loại hình kinh doanh mới được hình thành, trong đó có mua bán hàng trên mạng. Với hình thức này người tiêu dùng có thể tiếp cận với hàng hóa một cách dễ dàng và nhanh chóng hơn rất nhiều so với hình thức mua bán hàng truyền thống.

Hiện nay các hệ thống bán hàng trực tuyến đã tạo nhiều điều kiện thuận lợi để người mua có thể tiếp cận nhiều mặt hàng cùng một lúc. Tuy nhiên, một website thương mại thì luôn luôn mong muốn phát triển số lượng khách hàng, và muốn có nhiều khách hàng thì họ phải đa dạng hóa các loại sản phẩm để đáp ứng được nhu cầu mua sắm của nhiều loại khách hàng, do vậy số lượng sản phẩm và loại sản phẩm được trưng bày trong website ngày càng tăng và sẽ làm hạn chế khả năng giao tiếp chọn sản phẩm của khách hàng, họ phải duyệt qua nhiều liên kết, sàng lọc nhiều thông tin mới có thể tìm được sản phẩm mong muốn. Vậy làm sao hỗ trợ khách hàng trong công việc lựa chọn sản phẩm mua sắm? Cụ thể, những sản phẩm nào nên được đề xuất tiếp theo các sản phẩm đã được khách hàng đánh giá hoặc chọn trong giỏ hàng? Nên đề xuất bao nhiêu sản phẩm là tốt nhất cho khách hàng?

Để khách hàng có thể tìm và mua được một sản phẩm ưng ý thì một lời khuyên, một sự trợ giúp là rất quan trọng. Một người bán trong phương thức mua bán truyền thống là một lợi thế rất lớn. Do đó để hình thức mua bán qua mạng thực sự phát triển thì bên cạnh các lợi thế vốn có của mình việc có thêm một “người trợ giúp” là hết sức cần thiết. Hệ thống tư vấn được hình thành và phát triển không nằm ngoài mục đích đáp ứng những yêu cầu trên. Một hệ thống tư vấn tốt có thể đóng vai trò như người trung gian hỗ trợ khách hàng đưa ra quyết định chọn hàng. Tiện ích này đóng vai trò như một người bán hàng có khả năng thu thập thông tin về sở thích của khách hàng, sau đó tìm trong kho hàng vô tận của mình những mặt hàng thích hợp nhất với sở thích ñó. Thực chất của một hệ thống tư vấn này là quá trình hỗ trợ khách hàng đưa ra quyết định.

MỤC LỤC

**[Phần mở đầu 4](#_Toc7554)**

[1. Mục tiêu đề tài 4](#_Toc17548)

[2. Yêu cầu của đề tài 4](#_Toc2153)

[2.1. Yêu cầu 4](#_Toc17172)

[2.2. Các chức năng chính 4](#_Toc11458)

[2.3. Đối tượng và phạm vi nguyên cứu 5](#_Toc23898)

[3. Giải pháp thực hiện 5](#_Toc4759)

[3.1. Tìm hiểu thông tin 5](#_Toc11775)

[3.2. Môi trường phát triển và triển khai 6](#_Toc14817)

[4. Bố cục báo cáo 6](#_Toc3488)

**[Chương 1. Giới thiệu tổng quan 7](#_Toc14113)**

[1.1 Tìm hiểu về thương mại điện tử 7](#_Toc27540)

[1.2. Các ưu điểm của Website Thương mại điện tử 8](#_Toc7623)

[1.2.1. Thương mại điện tử giúp người bán 8](#_Toc12595)

[1.2.2. Thương mại điện tử giúp người mua 9](#_Toc28369)

[1.3. Tìm hiểu về hệ thống gợi ý 9](#_Toc11395)

[1.3.1 Hệ thống gợi ý là gì? 9](#_Toc13629)

[1.3.2. Lợi ích của hệ thống gợi ý 11](#_Toc9550)

**[Chương 2. Cài đặt trang web 13](#_Toc1869)**

[2.1. Thiết kế hệ thống 13](#_Toc6124)

[2.1.1 Kiến trúc hệ thống 13](#_Toc6124)

[2.1.2 Những công nghệ sử dụng để tạo website 13](#_Toc6124)

[2.2 Một số giao diện người dùng 16](#_Toc6829)

**[Chương 3. Áp dụng thuật toán luật kết hợp vào trang web 20](#_Toc11071)**

[3.1 Giới thiệu về thuật toán Apriori 20](#_Toc19190)

[3.2 Các bước của thuật toán Apriori . 21](#_Toc25529)

[3.3 Ví dụ thuật toán Apriori 22](#_Toc11561)

[3.4 Lý do chọn thuật toán Apriori để gợi ý cho website 25](#_Toc8997)

[3.5 Ứng dụng của thuật toán Apriori trong website thương mại điện tử 25](#_Toc31688)

[3.6 So sánh thuật toán Apriori với những thuật toán khác 27](#_Toc21420)

[3.6.1 Thuật toán Apriori 27](#_Toc12868)

[3.6.2 Thuật toán tìm kiếm mù 27](#_Toc163)

[3.6.3 Thuật toán tìm kiếm Best First Search 27](#_Toc16444)

**[Kết luận 28](#_Toc1470)**

[1. Kết quả đạt được 28](#_Toc23298)

[1.1. Ưu điểm 28](#_Toc4556)

[1.2. Nhược điểm 28](#_Toc4575)

[1.3. Cải thiện thuật toán luật kết hợp 28](#_Toc3006)

[2. Hướng phát triển 28](#_Toc14732)

[2.1 Ý nghĩa khoa học và thực tiễn của thuật toán luật kết hợp 29](#_Toc208)

[2.2 Kết luận 29](#_Toc4143)

PHẦN MỞ ĐẦU

## Mục tiêu đề tài

* Xây dựng một website mang lại thuận tiện cho người dùng, tiết kiệm thời gian tìm mua sản phẩm.
* Ứng dụng thuật toán luật kết hợp vào hệ thống gợi ý sản phẩm.

## Yêu cầu của đề tài

### Yêu cầu

#### Yêu cầu chung

- Tìm hiểu về hoạt động kinh doanh thương mại điện tử.

- Tìm hiểu về hệ thống gợi ý (Recommendation System).

- Xây dựng một trang web với đầy đủ các tính năng phục vụ cho việc bán hàng trực tuyến

- Xây dựng các chức năng giúp người quản trị dễ dàng quản lý nội dung trang web.

- Tìm hiểu thuật toán luật kết hợp Apriori.

##### **Xây dựng website thương mại điện tử.**

- Webiste thương mại điện tử có một số chức năng như sau:

+ Đối với khách hàng: Xem danh sách các sản phẩm, Xem thông tin chi tiết một sản phẩm, Tìm kiếm sản phẩm, Đưa sản phẩm vào giỏ hàng, Đặt mua, Thanh toán, Góp ý, Liên hệ, Xem các tin tức…

+ Đối với quản trị: Quản trị sản phẩm, quản trị nhân viên, Quản trị đơn hàng, Quản trị khách hàng, Quản trị tin tức, góp ý…

### Các chức năng chính

- Website bán thiết bị điện tử được xây dựng với tiêu chí phục vụ cho tất cả mọi người, trong lĩnh vực công nghệ, điện tử với mong muốn đáp ứng đầy đủ các mặt hàng, các chức năng của một website bán hàng trực tuyến.

- Website bao gồm các loại người dùng chính: Khách hàng viếng thăm, Khách hàng thành viên, Quản trị viên:

+ Đối với Khách hàng viếng thăm: Xem danh sách sản phẩm, Xem chi tiết sản phẩm, Đưa sản phẩm vào giỏ hàng, Cập nhật giỏ hàng, Đặt mua, Tìm kiếm, Đăng Ký…

+ Đối với Khách hàng thành viên: Ngoài những chức năng như khách hàng viếng thăm còn có thêm một số chức năng như Xem danh sách đơn hàng, Quản lý đơn hàng, Thanh Toán, Đăng nhập…

+ Đối với quản trị viên: Toàn quyền trong hệ thống và là người có quyền quyết định đối với khách hàng, đơn hàng, sản phẩm, tin tức, nội dung hiển thị.

### Đối tượng và phạm vi nguyên cứu

- Website bán thiết bị điện tử được xây dựng với tiêu chí phục vụ cho tất cả mọi người, trong lĩnh vực công nghệ, điện tử với mong muốn đáp ứng đầy đủ các mặt hàng, các chức năng của một website bán hàng trực tuyến.

- Với khoảng thời gian có hạn nên chúng em đã giới hạn phạm vi nguyên cứu ở một mức độ hợp lý cho mình:

+ Xây dựng các chức năng cho khách hàng.

+ Xây dựng các chức năng cho nhân viên quản trị.

+ Áp dụng thuật toán luật kết hợp Apriori vào hệ thống gợi ý.

## Giải pháp thực hiện

### Tìm hiểu thông tin

- Trong thời đại công nghiệp hóa - hiện đại hóa như hiện nay, Việt Nam đang trên đà phát triển, ngành công nghiệp điện, điện tử đang được quan tâm và phát triển mạnh. Các doanh nghiệp trong nước đua nhau sản xuất ra nhiều sản phẩm tiên tiến đáp ứng nhu cầu thị trường. Cùng với sự phát triển của nền kinh tế, chất lượng cuộc sống người dân được nâng cao. Nhu cầu tiêu dùng các mặt hàng điện tử cũng tăng mạnh theo, mang đến nhiều cơ hội đầu tư.

- Với sự nổi lên của Youtube, Amazon, Netflix và nhiều dịch vụ web khác, các hệ gợi ý ngày càng có nhiều vị trí hơn trong cuộc sống của chúng ta. Từ thương mại điện tử gợi ý cho người mua những bài báo mà họ có thể quan tâm đến quảng cáo trực tuyến gợi ý cho người dùng những nội dung phù hợp, phù hợp với sở thích của họ, hệ gợi ý ngày nay không thể tránh khỏi trong hành trình trực tuyến hàng ngày của chúng ta.Nói một cách tổng quát, hệ thống gợi ý là các thuật toán nhằm đề xuất các mặt hàng có liên quan cho người dùng (mặt hàng là phim để xem, tin nhắn văn bản để đọc, sản phẩm cần mua hoặc bất kỳ thứ gì khác tùy theo ngành).

- Hệ gợi ý thực sự rất quan trọng trong một số ngành vì chúng có thể tạo ra một khoản thu nhập khổng lồ khi chúng hoạt động hiệu quả hoặc cũng là một cách để nổi bật đáng kể so với các đối thủ cạnh tranh.

- Từ những thông tin tìm hiểu, yêu cầu đặt ra là: Để phục vụ tốt cho nhu cầu tìm kiếm các sản phẩm và gia tăng doanh số bán hàng thiết bị điện tử. Từ đó chúng em xây dựng website bán thiết bị điện tử áp dụng hệ thống gợi ý nhằm cung cấp các mặt hàng liên quan cho người dùng.

### Môi trường phát triển và triển khai

- Môi trường phát triển ứng dụng :

+ Công cụ xây dựng cơ sở dữ liệu SQL

+ Công cụ dùng để phân tích, thiết kế : Rational Rose Enterprise Edition

+ Công cụ đã dùng để xây dựng ứng dụng : Visual Studio Code

- Môi trường triển khai ứng dụng :

+ Hệ điều hành Windows 10

+ Để chạy được Website yêu cầu máy phải cài ít nhất một trình duyệt web.

+ Yêu cầu .NET FRAMEWORK 7.0 trở lên.

## Bố cục báo cáo

* Phần mở đầu
* Chương 1: Giới thiệu tổng quan
* Chương 2: Cài đặt trang web
* Chương 3: Áp dụng luật kết hợp vào trang web

CHƯƠNG 1 - GIỚI THIỆU TỔNG QUAN

## 1.1 Tìm hiểu về thương mại điện tử

- Công nghệ tiên tiến hiện nay giúp doanh nghiệp biến Website của mình thành những siêu thị hàng hóa trên Internet, biến người mua thực sự trở thành những người chủ với toàn quyền trong việc chọn lựa sản phẩm, tìm kiếm thông tin, so sánh giá cả, đặt mua hàng, ký kết hợp đồng với hệ thống tính toán tiền tự động, rõ ràng, trung thực.

- Thương mại điện tử (E-Commerce) là hình thái hoạt động kinh doanh bằng các phương pháp điện tử; là việc trao đổi "thông tin" kinh doanh thông qua các phương tiện công nghệ điện tử.

- Hiện nay có rất nhiều cách hiểu khác nhau về Thương mại điện tử. Nhiều người hiểu Thương mại điện tử là bán hàng trên mạng, trên Internet. Một số ý kiến khác lại cho rằng Thương mại điện tử là làm thương mại bằng điện tử. Những cách hiểu này đều đúng theo một góc độ nào đó nhưng chưa nói lên được phạm vi rộng lớn của Thương mại điện tử.

- Theo khái niệm này, Thương mại điện tử không chỉ là bán hàng trên mạng hay bán hàng trên Internet mà là hình thái hoạt động kinh doanh bằng các phương pháp điện tử. Hoạt động kinh doanh bao gồm tất cả các hoạt động trong kinh doanh như giao dịch, mua bán, thanh toán, đặt hàng, quảng cáo và kể cả giao hàng. Các phương pháp điện tử ở đây không chỉ có Internet mà bao gồm việc sử dụng các phương tiện công nghệ điện tử như điện thoại, máy FAX, truyền hình và mạng máy tính (trong đó có Internet). Thương mại điện tử cũng bao hàm cả việc trao đổi thông tin kinh doanh thông qua các phương tiện công nghệ điện tử. Thông tin ở đây không chỉ là những số liệu hay văn bản, tin tức mà nó gồm cả hình ảnh, âm thanh và phim video.

- Các phương tiện điện tử trong Thương mại điện tử

+ Điện thoại

+ Máy FAX

+ Truyền hình

+ Hệ thống thanh toán điện tử

+ Intranet / Extranet

- Mạng toàn cầu Internet / World Wide Web

- Các hình thức hoạt động Thương mại điện tử

+ Thư tín điện tử (E-mail)

+ Thanh toán điện tử

+ Trao đổi dữ liệu điện tử (EDI)

+ Trao đổi số hoá các dung liệu

+ Mua bán hàng hoá hữu hình

- Thương mại điện tử hiện nay đã trở nên khá quen thuộc và trở thành một môi trường thương mại không thể thiếu được trong đời sống kinh tế xã hội phát triển. Ở Việt Nam, Thương mại điện tử đang được các doanh nghiệp từng bước áp dụng vào hoạt động sản xuất kinh doanh của mình. Thương mại điện tử có những thế mạnh vượt trội mà không một loại hình kinh doành nào khác có được.

## 1.2. Các ưu điểm của Website Thương mại điện tử

- Ưu điểm tuyệt đối của Thương mại Điện tử là cho phép người sử dụng thực hiện các hoạt động kinh doanh ngay lập tức trên quy mô toàn cầu, từ việc quảng cáo công ty, tiếp thị sản phẩm, đàm phán và đặt hàng cho đến các khâu thanh toán, giữ liên hệ với khách hàng và hỗ trợ sau bán hàng. Bởi vì:

### 1.2.1. Thương mại điện tử giúp người bán

- Tiếp thị hiệu quả sản phẩm và dịch vụ của mình ra khắp thế giới

- Tạo kênh bán hàng trực tiếp tới khách hàng với quy mô rộng, tốc độ nhanh và chi phí giảm rất nhiều so với các kênh bán hàng truyền thống khác

- Mở ra khả năng xuất khẩu hàng ra nước ngoài

- Đơn giản hóa được các thủ tục hành chính, các công việc giấy tờ, tăng hiệu quả giao dịch thương mại

- Với Website Thương mại điện tử, doanh nghiệp tạo cho mình khả năng kinh doanh liên tục 24/24 giờ, liên tục 07 ngày trong tuần với chi phí rất thấp. Không cần nhân viên giám sát khách hàng như tại các siêu thị bình thường, không cần bỏ tiền thuê địa điểm bán hàng, không cần hệ thống kiểm tra, giới thiệu sản phẩm, không cần hệ thống tính tiền,... Tất cả đều được Website làm tự động, rất nhanh chóng và với độ chính xác tuyệt đối.

- Tại cùng 1 thời điểm, Website Thương mại điện tử có thể phục vụ hàng triệu lượt người mua hàng ở khắp nơi trên thế giới với các yêu cầu rất khác nhau về thông tin sản phẩm, chủng loại sản phẩm, giá cả, hình ảnh, chất lượng, mẫu mã,...

- Thông tin, giá cả sản phẩm được cập nhật, thay đổi một cách tức thời theo sự biến động của thị trường.

- Website Thương mại Điện tử đem lại khả năng kinh doanh mới cho doanh nghiệp: "Kinh doanh ngay cả khi bạn đang ngủ".

### 1.2.2. Thương mại điện tử giúp người mua

- Có thêm một hình thức mua hàng thuận tiện, dễ dàng, nhanh chóng.

- Có thêm một hình thức thanh toán mới tiện lợi, an toàn.

- Mở rộng sự chọn lựa khi mua hàng theo thị hiếu và nhu cầu.

- Có cơ hội mua sản phẩm và dịch vụ trực tiếp từ nhà sản xuất hoặc nhà cung cấp chính không qua trung gian.

- Người mua thực sự trở thành người chủ với toàn quyền lựa chọn sản phẩm, tìm kiếm bất kỳ thông tin nào về sản phẩm theo nhu cầu, so sánh giá cả, đặt mua hàng với hệ thống tính toán tiền tự động, đầy đủ, rõ ràng, trung thực và chính xác nhất.

## 1.3. Tìm hiểu về hệ thống gợi ý

### 1.3.1 Hệ thống gợi ý là gì?

- Trong tiếng anh hệ thống gợi ý được gọi là: Recommender Systems.

- Hệ thống gợi ý được xác định ở đây chính là kĩ thuật của trí tuệ nhân tạo, được nghiên cứu để cung cấp những gợi ý tự động tới người dùng hoặc khách hàng. Nó dựa trên dữ liệu về hành vi trong quá khứ của người dùng để hướng dẫn khách hàng định vị được những sản phẩm mà họ thích, giúp khách hàng có những quyết định tốt khi mua sắm online.

- Đối với thương mại điện tử, người mua và người bán không cần gặp gỡ trực tiếp, mà họ giao dịch với nhau trên trang web. Vì vậy, rất cần có một trợ lí bán hàng tự động. Việc tìm ra một “chuyên gia” tư vấn thông minh, thân thiện, am hiểu khách hàng trên các website vô cùng quan trọng và cũng chính là một thách thức. Vấn đề này có thể được giải quyết bằng việc tích hợp các kĩ thuật gợi ý (Recommender Systems) trong các website bán hàng. Hệ gợi ý như 1 chuyên gia tư vấn, dự đoán thông minh sở thích của khách hàng và cung cấp những thông tin mà họ thực sự quan tâm.

- Hệ thống đề xuất, hoặc hệ thống khuyến nghị (đôi khi thay thế ‘hệ thống’ bằng một từ đồng nghĩa như nền tảng hoặc động cơ), là một lớp con của hệ thống lọc thông tin tìm cách dự đoán “xếp hạng” hoặc “sở thích” mà người dùng sẽ đưa ra cho một mặt hàng .

- Hệ thống giới thiệu được sử dụng trong nhiều lĩnh vực khác nhau, với các ví dụ thường được công nhận dưới dạng trình tạo danh sách phát cho các dịch vụ video và âm nhạc, giới thiệu sản phẩm cho các cửa hàng trực tuyến hoặc giới thiệu nội dung cho các nền tảng truyền thông xã hội và giới thiệu nội dung web mở. Các hệ thống này có thể hoạt động bằng cách sử dụng một đầu vào duy nhất, như âm nhạc hoặc nhiều đầu vào trong và trên các nền tảng như tin tức, sách và truy vấn tìm kiếm. Ngoài ra còn có các hệ thống giới thiệu phổ biến cho các chủ đề cụ thể như nhà hàng và hẹn hò trực tuyến. Hệ thống đề xuất cũng đã được phát triển để khám phá các bài báo và chuyên gia nghiên cứu, cộng tác viên, và các dịch vụ tài chính.

- Các kĩ thuật hệ gợi ý đã và đang được nghiên cứu, ứng dụng một cách mạnh mẽ và mang lại lợi ích cho cả người cung cấp dịch vụ và người sử dụng dịch vụ.

- Hầu hết các trang thương mại lớn trên thế giới, như: Amazon.com, Alibaba.com, youtube.com, facebook.com, ebay.com, MovieFinder.com… đều sử dụng các kĩ thuật gợi ý trong website của mình để nâng cao trải nghiệm cho khách hàng, nâng cao chất lượng dịch vụ và thu lại lợi nhuận tốt hơn.

- Ở Việt Nam, những trang thương mại điện tử thành công, như: lazada.vn, Vatgia.com, Shopee.com, www.fptshop.com.vn, Thegioididong.com, Tiki.vn… đều có sử dụng hệ gợi ý trong hệ thống thông tin. Tuy nhiên, hiện nay, phần lớn trang thương mại điện tử trong nước chưa tích hợp các chức năng này, hoặc nếu có cũng chỉ dừng ở mức gợi ý chung chung, chưa có tính cá nhân hóa.

- Hệ thống gợi ý dần trở thành một thành phần không thể thiếu của các sản phẩm điện tử có nhiều người dùng. Các sản phẩm cá nhân hóa điện tử ngày càng phổ biến với mục đích mang sản phẩm phù hợp tới người dùng hoặc giúp người dùng có các trải nghiệm tốt hơn. Nếu quảng cáo sản phẩm tới đúng người dùng, khả năng các món hàng được mua nhiều hơn. Nếu gợi ý một video mà người dùng nhiều khả năng thích hoặc gợi ý kết bạn đến đúng đối tượng, họ sẽ ở lại trên nền tảng của bạn lâu hơn. Khi họ ở trên nền tảng của bạn lâu hơn, họ sẽ nhìn thấy nhiều quảng cáo hơn và lợi nhuận từ quảng cáo sẽ lại càng nhiều hơn.

### 1.3.2. Lợi ích của hệ thống gợi ý

- Các lợi ích chính của việc ứng dụng các kĩ thuật gợi ý cho thương mại điện tử (Jordan, 2016; Stephan, 2019):

– Hệ gợi ý giúp doanh nghiệp giảm thời gian tìm kiếm sản phẩm cho khách, tăng số lượt xem sản phẩm trên các gian hàng, tăng số lượng mua hàng của khách, tăng doanh số bán hàng.

– Hệ gợi ý làm tăng giá trị trung bình của đơn hàng (Average Order Value).

– Hệ gợi ý làm tăng tỉ lệ chuyển đổi của khách (CTR – Conversionrate of visitor): Theo nghiên cứu của Barilliance (Stephan, 2019) thấy rằng, tỉ lệ chuyển đổi mua hàng của khách click vào các gợi ý trên website tăng gấp 5,5 lần so với không click vào gợi ý.

– Hệ gợi ý cải thiện việc bỏ giỏ hàng lên tới 4,35% (Stephan, 2019).

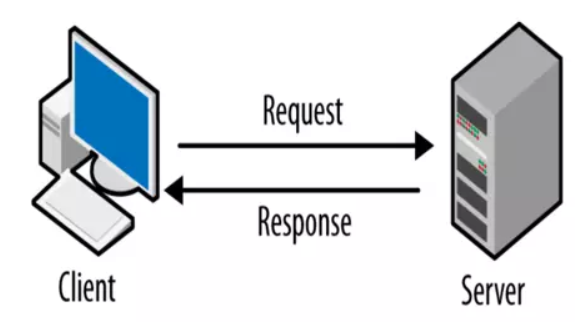
– Hệ gợi ý tăng doanh thu cho các doanh nghiệp. Cụ thể, mang lại 31% doanh thu trên các trang thương mại điện tử. Có đến 12% khách mua hàng đến từ các phương pháp gợi ý (Stephan, 2019).

CHƯƠNG 2 - CÀI ĐẶT TRANG WEB

## 2.1. Thiết kế hệ thống

### 2.1.1. Kiến trúc hệ thống

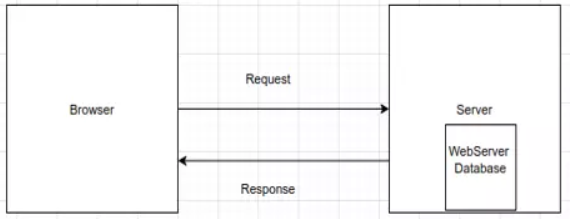
- Một hệ thống bất kì nào đó đều sẽ có 1 kiến trúc vật lý riêng và chúng đều được chia làm 2 phần chính là CLIENT và SERVER.  
- CLIENT tức là trình duyệt web.  
- SERVER (máy chủ) có thể hiểu như là 1 cái máy tính của bạn chứa RAM, CPU, ổ cứng... và các phần mềm phục vụ cho việc phát triển web.  
- Máy bình thường cũng có thể trở thành 1 máy chủ nếu nó được cài phần mềm Server Software và kết nối vs internet.



***a, Hệ thống nhỏ***

- Ví dụ: Tin tức, ứng dụng đồ án,các dự án chỉ truyền tải nội dung.…

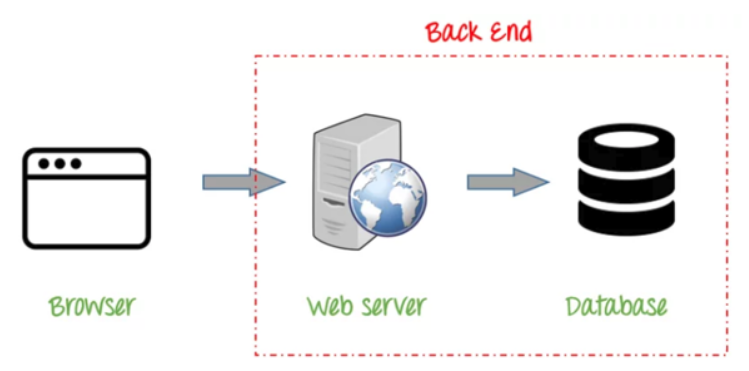
- Với hệ thống nhỏ thì tất cả sẽ làm trong 1 máy chủ (webserver, database server) tức là server chúng ta sẽ chứa và đảm nhiệm cũng như xử lý hết các công việc từ trình duyệt yêu cầu.



***b, Hệ thống trung bình***

-Ví dụ: dự án công ty, web bán hàng,..…

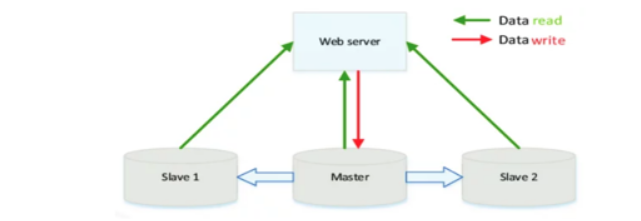
* WebServer: Đảm nhiệm xử lý yêu cầu bên server tức là nó sẽ sử dụng một ngôn ngữ (PHP, Java,...) tương thích với nó thường sẽ có 1 framework (Laravel,...) để xử lý nghiệp vụ và logic.
* Database Server: Dùng để chứa dữ liệu.



***c, Hệ thống lớn***

- Ví dụ: Facebook, tiki, các dự án có số người truy cập lớn.

- Người ta sử dụng kiến trúc: **Kiến trúc master/slave**



- Sharding database có thể hiểu đơn giản là chia nhỏ các database ra theo một số đặc điểm chung  
- Ví dụ: Người việt nam sẽ truy cập đến database ở việt nam, mỹ sẽ truy cập database ở mỹ.

### 2.1.2. Những công nghệ sử dụng để tạo website

***a, Yêu cầu chung về công nghệ:*** Việc lựa chọn công nghệ để xây dựng website phải đáp ứng các yêu cầu chung sau đây.

* Giải pháp công nghệ phải mang lại sự thành công của dự án, điều đó có nghĩa là hệ thống phải hoạt động tốt, đáp ứng mục đích đề ra.
* Giải pháp công nghệ lựa chọn phải đảm bảo hệ thống hoạt động với tính ổn định , lien tục, lâu dài và tính bảo mật cao.
* Đáp ứng khả năng tăng trưởng dữ liệu và thay đổi mở rộng nâng cấp website trong tương lai.
* Chi phí bản quyền ít tốn kém.

***b, Công nghệ Thiết kế webite:*** Bọn em đề xuất sử dụng công nghệ dựa trên nền tảng mã nguồn mở: Ngôn ngữ phát triển: PHP phiên bản 8.0.0

- Lý do:

* Các chuyên gia công nghệ khẳng định sử đơn giản, tin cậy và hiệu quả là ba ưu tiên hàng đầu của ngôn ngữ PHP – ngôn ngữ lệnh cho phép chuyển thông tin từ cơ sở dữ liệu tới máy chủ web cùng nhiều chức năng khác.
* PHP là ngôn ngữ phát triển web phổ biết nhất hiện nay, có đến hơn 80% website trên Thế giới được phát triển với ngôn ngữ PHP mà điển hình là 01 trang web lớn nhất Thế giới hiện nay là Google và Yahoo!.
* Các ứng dụng viết bằng ngôn ngữ PHP chạy tốt trên tất cả các Hệ điều hành hiện nay và tương tác với nhiều hệ quản trị CSDL khác nhau. Các ứng dụng viết bằng ngôn ngữ PHP chạy ổn định, đáp ứng lượng truy cập lớn, dễ bảo trì, dễ nâng cấp và mở rộng khi có nhu cầu.
* Ngoài ra ngôn ngữ PHP còn có ưu điểm hơn ASP về mặt kinh tế. Khi Việt Nam đã gia nhập WTO, vấn đề bản quyền phần mềm là một trong những cú “Shock” mà các tổ chức đơn vị sử dụng công nghệ có bản quyền gặp phải. Trong đó, các website sử dụng công nghệ ASP và hệ quản trị CSDL Micrsoft Server là một ví dụ điển hình.

***c, Hệ quản trị CSDL:*** MySQL   
- Lý do:

* Với lựa chọn công nghệ mã nguồn mở PHP, thì hệ quản trị CSDL được lựa chọn là MySQL. Hệ quản trị này hoạt động ổn định, đáp ứng lượng truy cập lớn và hỗ trợ tốt Unicode.
* Chương trình quản lý Database nguồn mở này được hai nhà pht triển người Thụy Điển xây dựng vào những năm 90 và đã thu hút ngay được hơn 5 triệu người sử dụng trên toàn cầu.

***d, Web server:*** Phần mềm máy chủ Web Apache.  
- Lý do:

* Theo Netcraft, công ty dịch vụ Internet hàng đầu ở Anh, gần 50 triệu trang web trên Thế giới hiện nay sử dụng ứng dụng máy chủ Web Apache, chiếm hơn 70% lĩnh vực Web server. Đối thủ lớn nhất của Apache là Internet Information Services (IIS) của Microsoft cũng chỉ có 20%.
* Chuyên gia Rich Morrow nhận định: “Điều khiến Apache trở nên phổ biến là nó có tốc độ nhanh và dễ tùy biến”.

***e, Front end:***

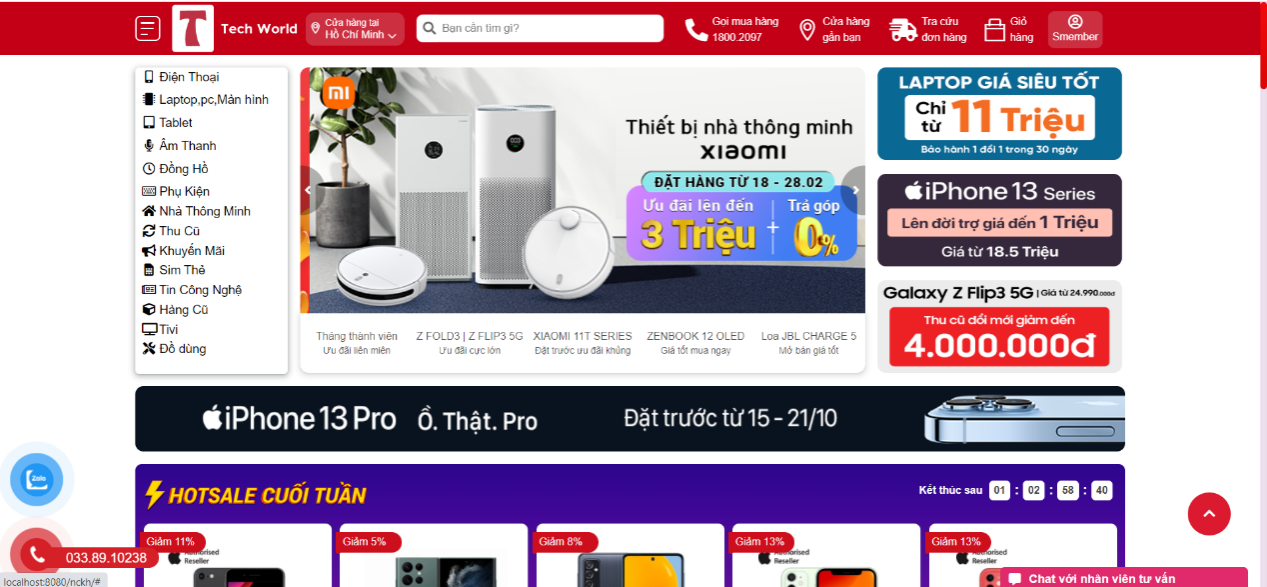
- Bootstrap 5.0

- Jquery 3.6.2

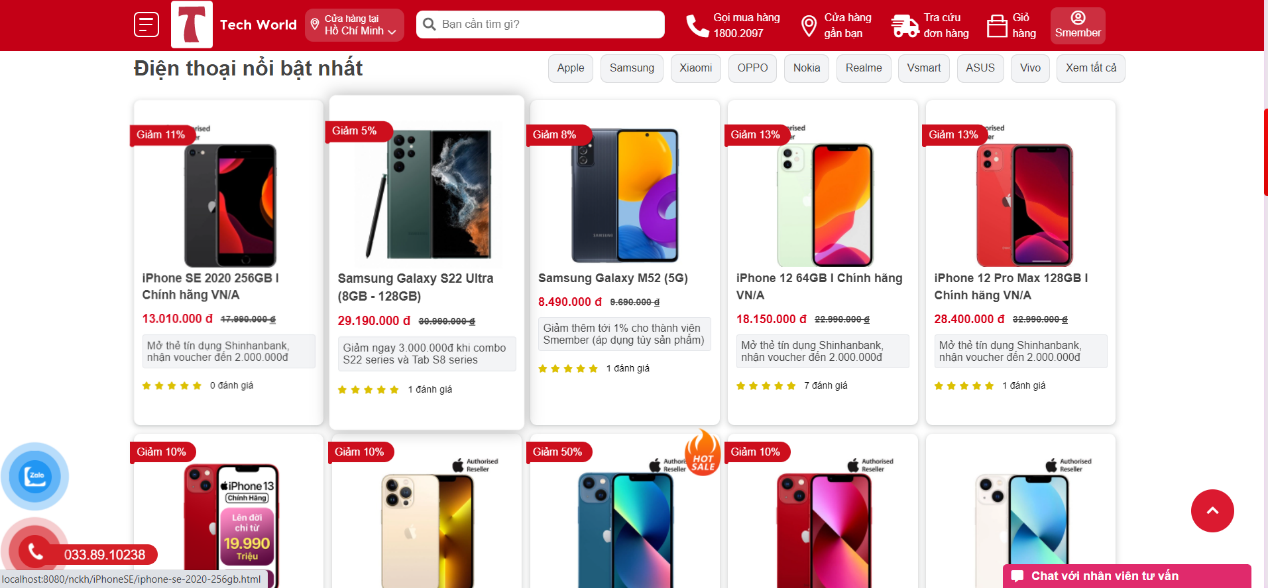
*-* Bootstrap và Jquery là 2 thư viện phổ biến và dễ sử dụng có thể tạo ra 1 giao diện web nhanh chóng và dễ dàng giúp tiết kiệm rất nhiều chi phí , thời gian và dễ bảo trì

## 2.2 Một số giao diện người dùng

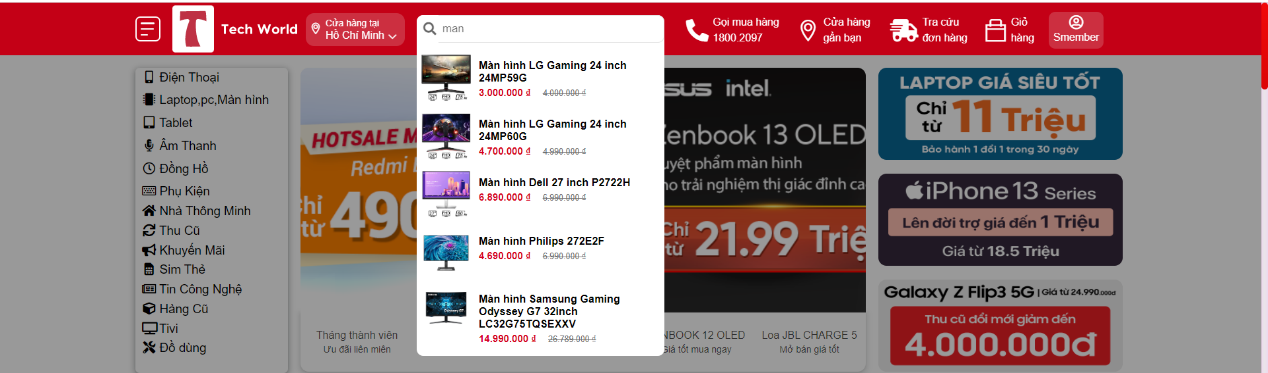
### 2.2.1 Giao diện khách hàng



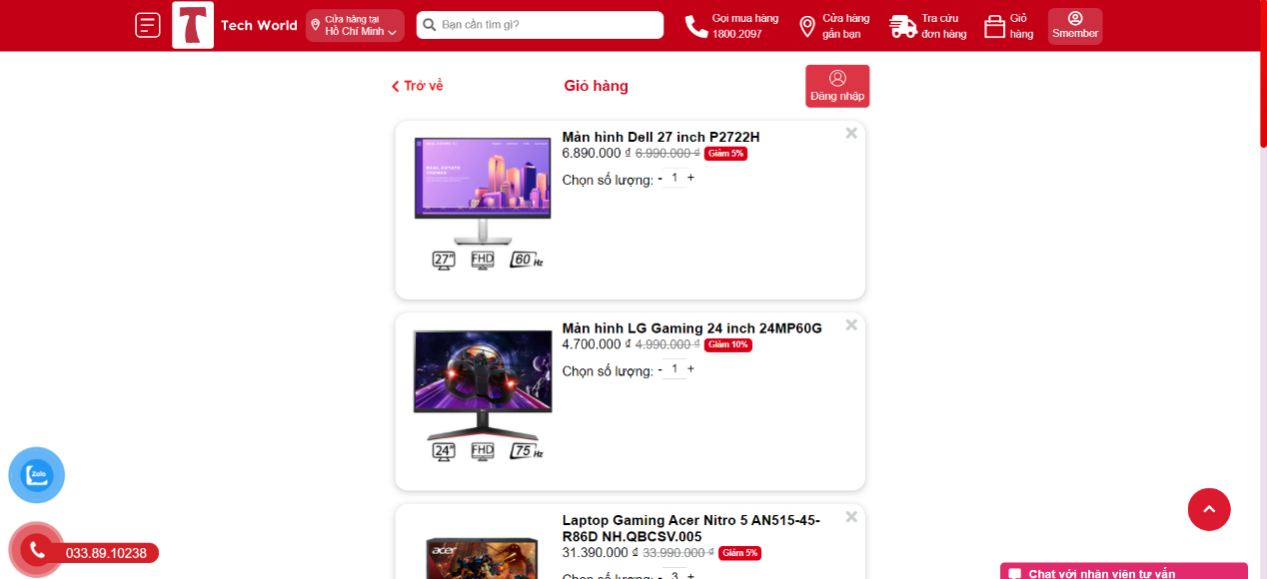
*Hình 2.3.1 : Ảnh trang chủ*



*Hình 2.3.2 : Ảnh sản phẩm của trang chủ*

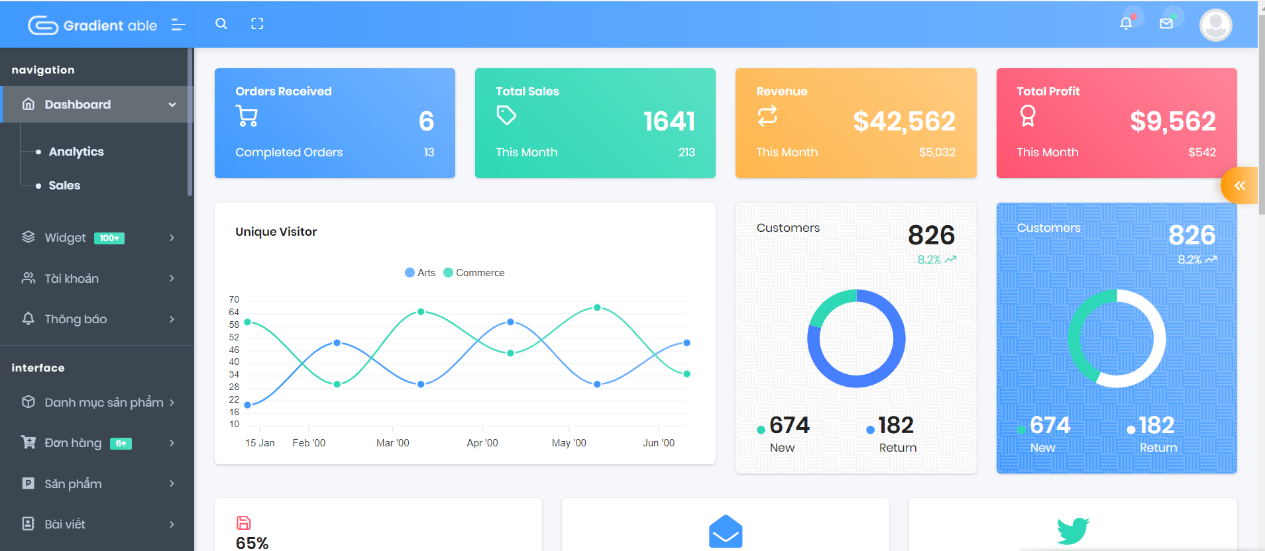


*Hình 2.3.3 : Ảnh tìm kiếm sử dụng api*

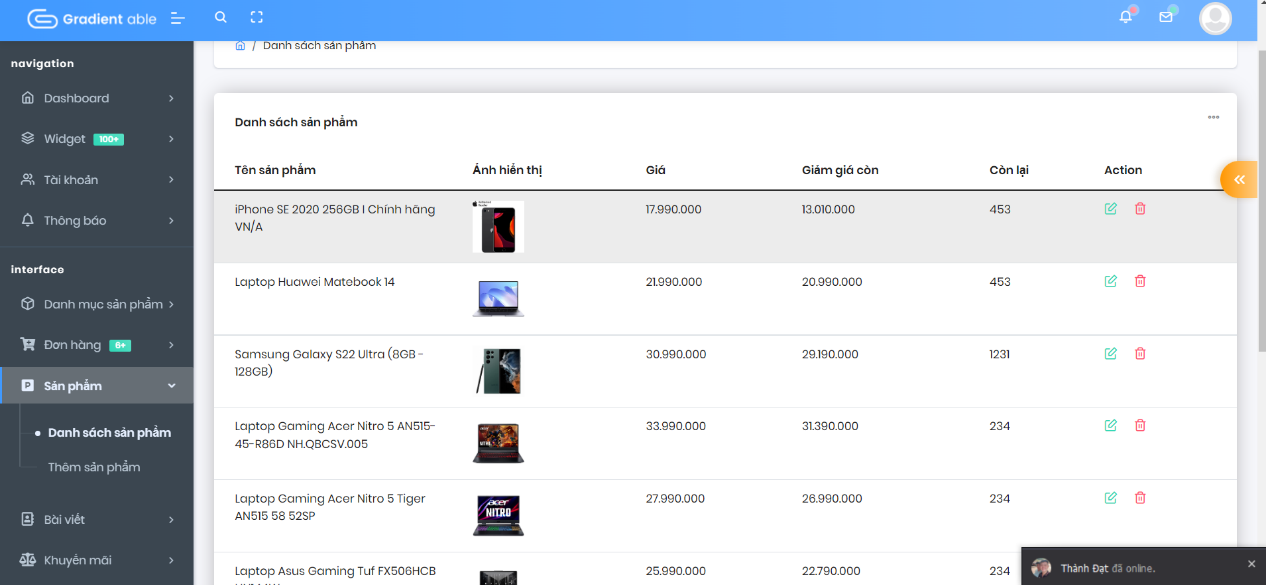


*Hình 2.3.4 : Ảnhgiỏ hàng*

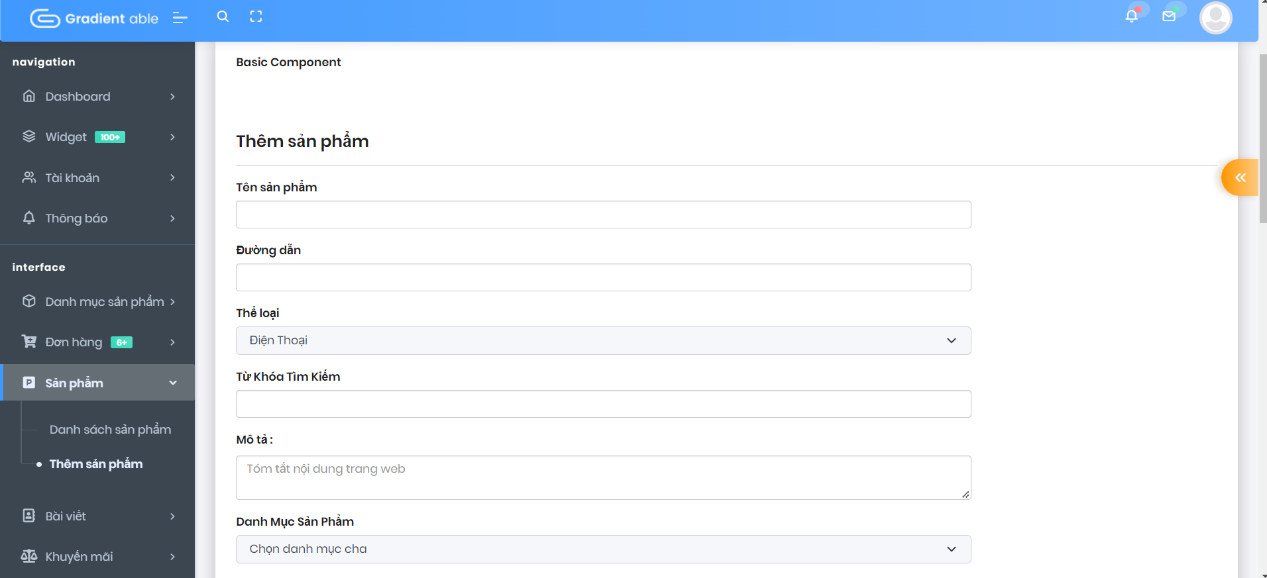
### 2.2.2. Giao diện quản trị viên



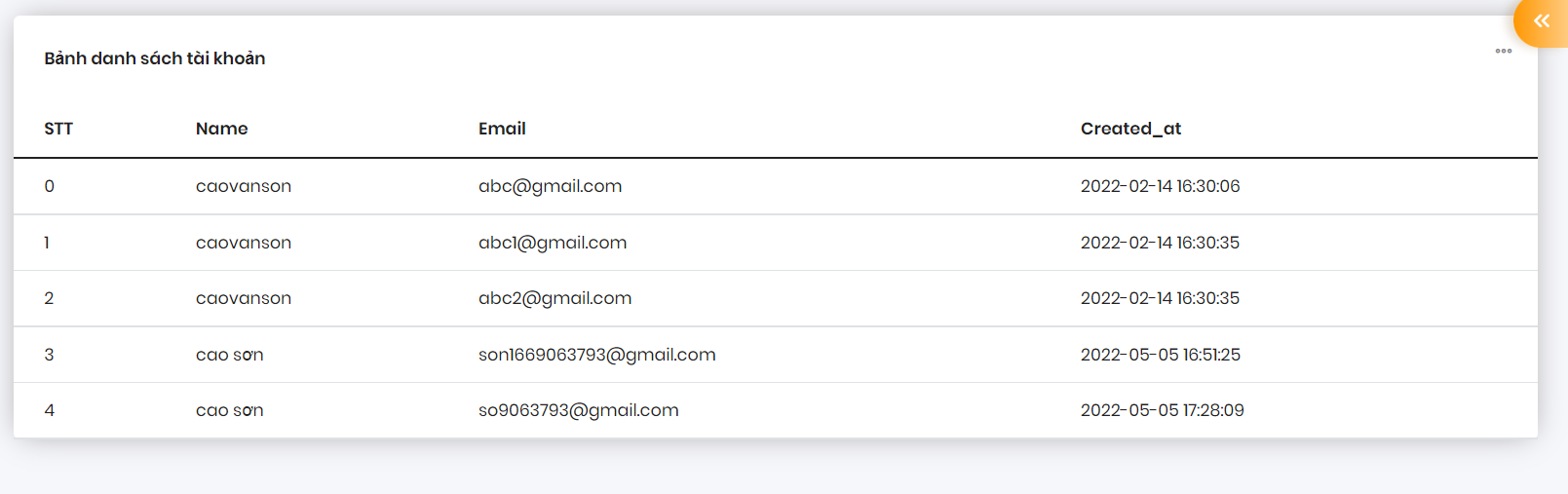
*Hình 2.3.5 : Ảnh trang home*



*Hình 2.3.6 : Ảnh bảng danh sách sản phẩm*



*Hình 2.3.7 : Ảnh bảng thêm sản phẩm*



*Hình 2.3.8 : Ảnh bảng tài khoản người dùng*

CHƯƠNG 3 - ÁP DỤNG THUẬT TOÁN LUẬT KẾT HỢP VÀO TRANG WEB

## Giới thiệu về thuật toán Apriori

 - Apriori là thuật toán khả sinh được đề xuất bởi R. Agrawal và R. Srikant vào năm 1993 để khai thác các tập item đối với các luật kết hợp kiểu bool. Tên của thuật toán dựa trên việc thuật toán sử dụng *tri thức trước* (*prior knowledge*) của các thuộc tính tập item phổ biến, chúng ta sẽ thấy sau đây. Apriori dùng cách tiếp cận lặp được biết đến như tìm kiếm *level-wise*, với các tập *k* item được dùng để thăm dò các tập (*k*+1) item. Đầu tiên, tập các tập 1 item phổ biến được tìm thấy bằng cách quét cơ sở dữ liệu để đếm số lượng từng item, và thu thập những item thỏa mãn độ hỗ trợ tối thiểu. Tập kết quả đặt là *L*1. Tiếp theo, *L*1được dùng để tìm *L*2, tập các tập 2 item phổ biến, nó được dùng để tìm*L*3, và cứ thế tiếp tục, cho tới khi tập *k* item phổ biến không thể tìm thấy. Việc tìm kiếm cho mỗi *L*kđòi hỏi một lần quét toàn bộ cơ sở dữ liệu.

- Trước khi đi vào chi tiết của thuật toán Apriori đầu tiên chúng ta sẽ tìm hiểu xác định một vài thuật ngữ phổ biến được sử dụng trong thuật toán.

* Itemset- Itemset là tập hợp của những item trong cơ sở dữ liệu mà nó được xác định bởi I={i1, i2,…. in}, trong đó n là số lượng item.
* Transaction- Transaction là một thành phần cơ sở dữ liệu mà nó bao gồm tập hợp các item. Transaction được ký hiệu là T và T  I. Một Transaction chứa tâp hợp các item T={i1, i2,…. in}.
* Minimum support – Minimum support là điều kiện cần được đáp ứng bởi các item đề ra để có thể tiến hành xử lý item kế tiếp có thể. Minimum support có thể được xem như là một điều kiện giúp loại bỏ các tâp không phổ biến trong bất kỳ cơ sở dữ liệu. Thường sử dụng Minimum support cho mô hình tỷ lệ phần trăm.
* Frequent itemset (tập phổ biến) -  các Itemset đáp ứng các tiêu chí điều kiện minimum support thì được gọi là tập phổ biến. Nó được ký hiệu là Li trong đó i chỉ   i-itemset.
* Candidate itemset (ứng viên tập phổ biến) - ứng viên tập phổ biến là các item chỉ được xem xét xử lý. Ứng viên tập phổ biến là tất cả các kết hợp có thể có của tập phổ biến. Nó thường được ký hiệu Ci trong đó I chỉ i-itemset.
* Support – Độ hữu dụng của một luật có thể được đo với sự giúp đỡ của ngưỡng hỗ trợ. Support giúp chúng ta đo như thế nào các giao tác có  tập phổ biến mà nó phù hợp với ý nghĩa cả hai phía cạnh trong luật kết hợp. Xem xét hai item A và B. Để tính toán support của A->B theo công thức như sau:  
                
* Confidence – Confidence chỉ sự chắc chắn của các luật. Thông số này cho phép chúng ta đếm mức độ thường xuyên một giao tác của tập phổ biến phù hợp với ý nghĩa cả phía cạnh bên trái với phía cạnh bên phải. các tập phổ biến không đáp ứng các điều kiện trên có thể được loại bỏ. Xem xét hai item A và B. Để tính toán confidence của A->B theo công thức sau:  
                  
  **Chú ý:** Conf(A->B) có thể không bằng conf(B->A)

Để tăng hiệu quả của việc phát sinh level-wise của tập item phổ biến, một tính chất quan trọng gọi là **tính chất Apriori**(**Apriori property**), được giới thiệu dưới đây, dùng để giảm không gian tìm kiếm. Chúng ta sẽ mô tả tính chất này trước, rồi mới xem một ví dụ minh họa cách sử dụng nó.

**Tính chất Apriori** : *Tất cả các tập con không rỗng của một tập item phổ biến cũng phải là phổ biến.*

Tính chất Apriori này dựa theo nhận xét sau. Theo định nghĩa, nếu một tập item *I* không thỏa ngưỡng độ hỗ trợ tối thiểu, *min\_sup*, thì *I* không là phổ biến, do đó, *P(I)< min\_sup*. Nếu một item *A* được thêm vào tập item *I*, thì tập item được tạo thành (vd, *I  A*) không thể xuất hiện thường xuyên hơn *I*. Do đó, *I  A* cũng không phổ biến; do đó, P(I ** A)*< min\_sup*.

Tính chất này thuộc về loại đặc biệt của các thuộc tính gọi là **chống đơn điệu**(**antimonotone**) trong nghĩa rằng *nếu một tập không thể qua một cuộc kiểm tra, tất cả các tập cha (superset) của nó cũng sẽ thất bại với một cuộc kiểm tra tương tự*. Đó gọi là *chống đơn điệu* là vì thuộc tính này là đơn điệu (monotonic) trong ngữ cảnh của việc thất bại một cuộc kiểm tra.

## 3.2 Các bước của thuật toán Apriori

## - Apriori dùng cách tiếp cận lặp được biết đến như tìm kiếm *level-wise*, với các tập *k* item được dùng để thăm dò các tập (*k*+1) item.

1. Đầu tiên, tập(frequent 1- itemsets) phổ biến 1 được tìm thấy ký hiệu là C1
2. Bước tiếp theo là tính support có nghĩa là sự xuất hiện của các item trong cơ sở dữ liệu. Điều này đòi hỏi phải duyệt qua toàn bộ cơ sở dữ liệu.
3. Sau đó, bước cắt tỉa được thực hiện trên C1 trong đó những item được so sánh với thông số minimum support. Những item thỏa điều kiện minimum support mới được xem xét cho tiến trình tiếp theo ký hiệu là L1.
4. Sau đó, bước phát sinh các bộ ứng viên được thực hiện trong đó tập phổ biến 2 được tạo ra ký hiệu là C2.
5. Một lần nữa, cở sở dữ liệu được duyệt để tính toán support của 2 tập phổ biến. Theo minimum support, các bộ ứng viên tạo ra được kiểm tra và chỉ những tập phổ biến nào thỏa điều kiện minimum support thì tiếp tục được sử dụng tạo ra bô ứng viên tập phổ biến 3.

- Bước trên tiếp tục cho đến khi không có tập phổ biến hoặc bộ ứng viên có thể được tạo ra.

## 3.3 Ví dụ thuật toán Apriori

Bảng 1 biểu diễn một giao dịch cơ sở dữ liệu có 4 giao dịch.

TID là một nhận dạng duy nhất cho mỗi giao dịch.

TABLE\_1

|  |  |
| --- | --- |
| TID | Items |
| T001 | A,C,D |
| T002 | B,C,E |
| T003 | A,B,C,E |
| T004 | B,E |

Thực hiện bước đầu tiên là chức năng duyệt cơ sở dữ liệu để xác định số lượng sự xuất hiện cho một item cụ thể. Sau bước đầu tiên chúng ta sẽ có  C1 được  thể hiện trong Table 2.

TABLE\_2 C1

|  |  |
| --- | --- |
| Itemset | Support |
| {A} | 2 |
| {B} | 3 |
| {C} | 3 |
| {D} | 1 |
| {E} | 3 |

Bước tiếp theo là bước cắt tỉa, trong đó support tập phổ biến được so sánh với minimum support. Các tập phổ biến thỏa măn minimum support sẽ được xử lý tiếp tục. Giả sử rằng minimum support là 2. Chúng ta sẽ có được L1 từ bước này. TABLE 3 cho thấy kết quả cắt tỉa.

TABLE\_3 L1

|  |  |
| --- | --- |
| Itemset | Support |
| {A} | 2 |
| {B} | 3 |
| {C} | 3 |
| {E} | 3 |

 Bây giờ bước tiếp theo phát sinh ứng viên được thực hiện trong đó tất cả ứng viên có thể có 2 tập phổ biến ứng viên được tạo. Bảng này được ký hiệu là C2. TABLE 4 cho thấy tất cả khả năng kết hợp mà có thể tạo ra từ TABLE 3 tập phổ biến.

TABLE\_4 C2

|  |  |
| --- | --- |
| Itemset | Support |
| {A,B} | 1 |
| {A,C} | 2 |
| {A,E} | 1 |
| {B,C} | 2 |
| {B,E} | 3 |
| {C,E} | 2 |

Bây giờ cắt tỉa phải được thực hiện trên cơ sở các điều kiện minimum support. Từ TABLE 4 hai tập phổ biến sẽ được loại bỏ. Sau khi cắt tỉa chúng ta nhận được kết quả như sau:

TABLE\_5  L2

|  |  |
| --- | --- |
| Itemset | Support |
| {A,C} | 2 |
| {B,C} | 2 |
| {B,E} | 3 |
| {C,E} | 2 |

Các quá trình tương tự được tiếp tục cho đến khi không có tập phổ biến hoặc bộ ứng viên có thể tạo ra. Tiến trình được mô tả trong TABLE 6 và TABLE 7.

TABLE\_6  C3

|  |  |
| --- | --- |
| Itemset | Support |
| {A,B,C} | 1 |
| {A,B,E} | 1 |
| {B,C,E} | 2 |

TABLE\_7 L3  (Kết quả cuối cùng)

|  |  |
| --- | --- |
| Itemset | Support |
| {A,B,C} | 1 |
| {A,B,E} | 1 |
| {B,C,E} | 2 |

Mã giải thuật toán Apriori\_Algorithm()

Apriori\_Algorithm()

{  
Ck: Candidate itemset of size k  
Lk : frequent itemset of size k  
L1 = {frequent items};  
**for** (k = 1; Lk!=0; k++)  
**{**  
          Ck+1 = candidates generated from Lk;  
**foreach** transaction t in database do  
                 increment the count of all candidates in  Ck+1                       
                    that are contained in t  
          Lk+1  = candidates in Ck+1 with min\_support  
 **}**  
return kLk;  
}

**3.4 Lý do chọn thuật toán Apriori để gợi ý cho website**

+ Trong lĩnh vực Data Mining, mục đích của luật kết hợp (Association Rule - AR) là tìm ra các mối quan hệ giữa các đối tượng trong khối lượng lớn dữ liệu.

+ Dễ dàng cập nhật các thông tin cần thiết như: hồ sơ học sinh, theo dõi nề nếp, điểm học tập, ...

+ Giúp thuận lợi cho người sử dụng khai thác dữ liệu theo chủ đề với các nguồn và khoảng thời gian khác nhau.

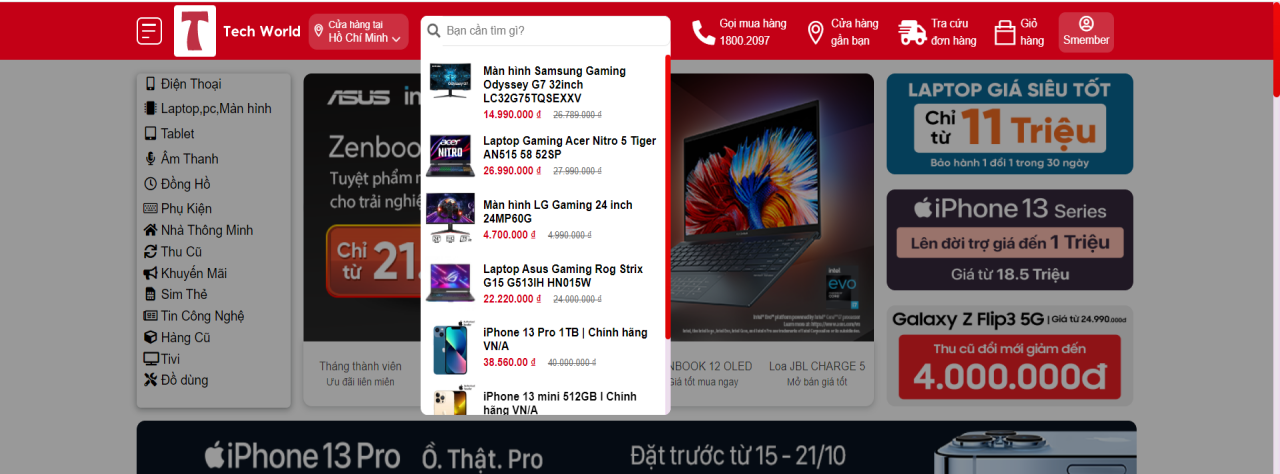
+ Dễ dàng tạo ra các báo cáo đơn giản, phù hợp với nhiều người sử dụng ở nhiều trình độ khác nhau.

+ Hỗ trợ việc xây dựng kho dữ liệu.

+ Thiết kế, cập nhật hệ thống giúp xử lý dữ liệu nhanh. Dẽ dàng tích hợp dữ liệu tác nghiệp mới và tạo các báo cáo mới theo yê cầu của người dùng.

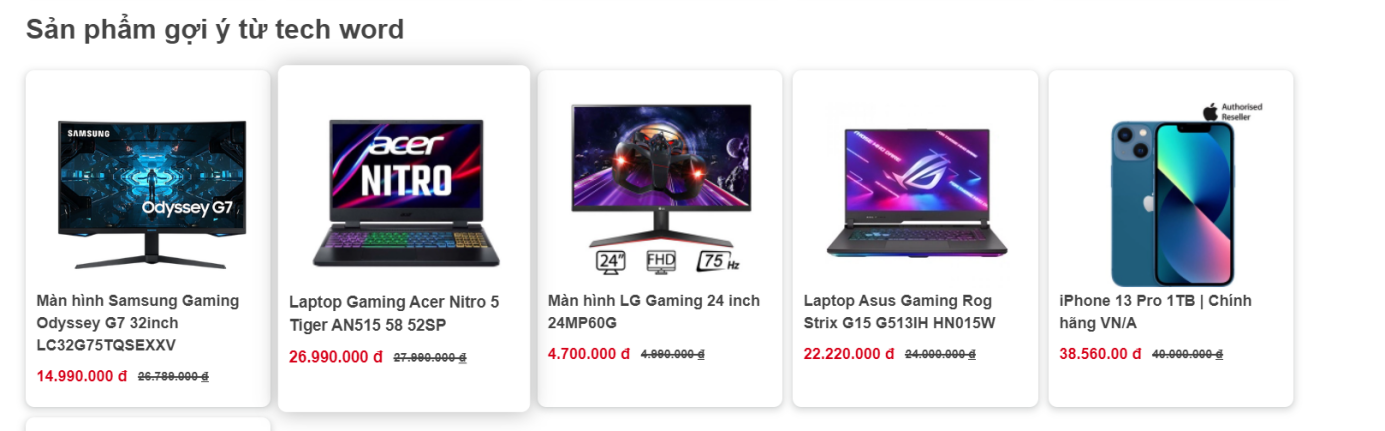
## 3.5 Ứng dụng của thuật toán Apriori trong website thương mại điện tử

**+ Công cụ tìm kiếm :** Chúng là những trang web được sử dụng nhiều nhất và được truy cập nhiều nhất hiện nay, những trang như Google và Bing, chúng là những trang mà mọi người đều truy cập khi mở trình duyệt. Và với điều này, các trang web này là tốt nhất cho tiếp thị trang web thương mại điện tử. Và sau đây chúng em sẽ giới thiệu cho bạn các chiến lược khác nhau để quảng cáo trang thương mại điện tử trên các trang web tìm kiếm này.



Hình 3.4.1 : Gợi ý tìm kiếm

**+ Gợi ý tìm kiếm :** Do 1 phần thuật toán phải duyệt dữ liệu rất nhiều lần để đưa ra kết quả chính xác lên bọn em quyết định để web lấy api data ngầm và đưa ra kết quả khi người dùng nhấn vào thanh công cụ tìm kiếm luôn . Data được lấy từ database để sử lý chỉ lấy khoảng 100 dữ liệu đầu tiên mới nhất để xử lý vì nếu lấy nhiều quá thì thuật toán sẽ xử lý rất lâu và có khả năng sẽ tràn bộ nhớ ảnh hưởng tới chất lượng người dùng.



Hình 3.4.2 : Gợi ý tìm kiếm khi người dùng mua sản phẩm

**+ Gợi ý tìm kiếm khi người dùng mua sản phẩm :** Tương tự như cách trên dữ liệu sẽ được lấy qua API và để cho thuật toán xử lý ngầm , các sản phẩm đưuọc gợi ý ra sẽ hiện thành 1 danh sách những sản phẩm có độ tin tưởng cao nhất.

## 3.6 So sánh thuật toán Apriori với những thuật toán khác

### 3.6.1 Thuật toán Apriori

+ Thuật toán Apriori đơn giản, dễ hiểu dễ cài đặt.

+ Nhờ các tính chất Apriori để cắt tỉa được nhiều nhánh trên giàn(lattice), giảm bớt đáng kế việc sinh các tập mục ứng viên và kiếm tra tính phố biến của các tập ứng viên đó.

=> Apriori dựa vào những dữ liệu ở quá khứ để có thể dự đoán chuẩn về tương lai và hiện tại . Dữ liệu càng ít thì độ tin tưởng càng cao .

### 3.6.2 Thuật toán tìm kiếm mù

+ Là kỹ thuật tìm kiếm mà trong đó chúng ta không có hiểu biết gì về các đối tượng để có hướng dẫn tìm kiếm mà chỉ đơn thuần xem xét các đối tượng theo một hệ thống nào đó để phát hiện ra đối tượng cần tìm.

=> Kết quả cũng chỉ dựa vào ăn may , độ tin tưởng thấp

### 3.6.3 Thuật toán tìm kiếm Best First Search

+ Tìm kiếm tốt nhất - đầu tiên (Best First Search) là tìm kiếm theo bề rộng (Breadth First Search) được hướng dẫn bởi hàm đánh giá. Tư tưởng của thuật toán này là việc tìm kiếm bắt đầu tại nút gốc và tiếp tục bằng cách duyệt các nút tiếp theo có giá trị của hàm đánh giá là thấp nhất so với các nút còn lại nằm trong hàng đợi.

+ Tìm kiếm kinh nghiệm, sử dụng hàm đánh giá

+ Một hàm đánh giá tốt có thể giảm thời gian và không gian nhớ một cách đáng kể.

=> Quá trình tìm kiếm có thể đi xa khỏi lời giải. Kỹ thuật này chỉ xét một phần của không gian và coi đó là phần hứa hẹn hơn cả. Có thể bị kẹt trong vòng lặp.

KẾT LUẬN

## 1. Kết quả đạt được

### Ưu điểm

* Thuật toán Apriori đơn giản, dễ hiểu dễ cài đặt.
* Nhờ các tính chất Apriori để cắt tỉa được nhiều nhánh trên giàn(lattice), giảm bớt đáng kế việc sinh các tập mục ứng viên và kiếm tra tính phố biến của các tập ứng viên đó.

### Nhược điểm

* Nhờ các tính chất Apriori để cắt tỉa được nhiều nhánh trên giàn , giảm bớt đáng kế việc sinh các tập mục ứng viên và kiếm tra tính phố biến của các tập ứng viên đó.
* Thực hiện việc tính độ phổ biến nhiều, đơn điệu.

### Cải thiện thuật toán luật kết hợp

* Giảm số lần duyệt CSDL
* Giảm số lượng tập ứng viên
* Qui trình tính độ phổ biến thuận tiện hơn

## 2. Hướng phát triển

- Việc thuật toán Apriori có thể làm là nhìn vào quá khứ và khẳng định rằng nếu một việc gì đó xảy ra thì sẽ có tỉ lệ bao nhiêu phần trăm sự việc tiếp theo sẽ xảy ra. Nó giống như nhìn vào quá khứ để dự đoán tương lại vậy, và việc này rất có ít cho các nhà kinh doanh. Ví dụ một siêu thị muốn nghĩ cách sắp xếp các gian hàng một cách hợp lí nhất, họ có thể nhìn vào lịch sử mua hàng và sắp sếp các tập sản phẩm thường được mua cùng nhau vào một chỗ. Hoặc một trang web tin tức muốn giới thiệu cho người dùng các bài viết liên quan đến nhau nhất, cũng có thể áp dụng quy luật tương tự.

- Việc thu thập và lưu trữ các kho chứa dữ liệu khổng lồ dẫn tới một yêu cầu cấp thiết là cần có những kỹ thuật và công cụ mới để tự động chuyển đổi lượng dữ liệu khổng lồ thành các tri thức có ích. Do vậy, khai phá dữ liệu (KPDL) nhằm phát hiện các tri thức mới giúp ích cho hoạt động của con người đã trở thành một lĩnh vực quan trọng của ngành Công nghệ thông tin.

- Khai phá dữ liệu đang được áp dụng rộng rãi trong nhiều lĩnh vực kinh doanh và đời sống khác nhau: marketing, tài chính–ngân hàng, bảo hiềm, khoa học,ytế, an ninh, giáo dục, internet ... rất nhiều công ty, xí nghiệp, các cơ quan đã áp dụng kỹ thuật khai phá dữ liệu vào hoạt động của mình và đã thu được nhiều thành tựu to lớn.

### 2.1 Ý nghĩa khoa học và thực tiễn của thuật toán luật kết hợp

**a, Về mặt lý thuyết**

* Phương pháp này được nhiều nhà khoa học nghiên cứu và có nhiều đóng góp trong thực tiễn.
* Có thể coi là tài liệu tham khảo về luật kết hợp.
* Người khác có thể tham khảo phát triển các ứng dụng riêng của mình.
* Ứng dụng Tin học vào quản lý thiết bị trong thương mại điện tử , công nghệ thông tin.

**b, Về mặt thực tiễn**

* Bảo hiểm, tài chính và thị trường chứng khoản: Phân tích tìnhhình tài chính và dự báo giá của các loại cổ phiếu trong thịtrường chứng khoán. Danh mục vốn và giá, lãi suất, dữ liệu thẻ tín dụng, phát hiện gian lận.
* Thống kê, phân tich dữ liệu và hỗ trợ ra quyết định.
* Điều trị y học và chăm sóc y tế: một số thông tin về chẩn đoánbệnh lưu trong các hệ thống quản lý bệnh viện. Phân tích mốiliên hệ giữa triệu chứng bệnh, chuẩn đoán và phương pháp điều trị (chế độ dinh dưỡng ...)
* Sản xuất và chế biến: quy trình, phương pháp chế biến và xử lý sự cố.
* Text mining&Web mining: phân lớp văn bản và các trang web, tóm tắt văn bản.
* Lĩnh vực khoa học: Quan sát thiên văn, dữ liệu gene, dữ liệu sinh vật học, tìm kiếm, so sánh các hệ gene và thông tin di truyền, mối liên hệ gene và một số bệnh di truyền…

### 2.2 Kết luận

Trên đây bọn em đã nghiên cứu về khái niệm luật kết hợp trong khai phá dữ liệu để tìm được mối liên hệ giữa các mục dữ liệu (items) của cơ sở dữ liệu bằng thuật toán Apriori . Trong thương mại điện tử tìm ra mối liên kết giữa sản phẩm và người dùng để làm cơ sở đánh giá đáng tin cậy.

TÀI LIỆU THAM KHẢO

[1] <https://www.levandong.com/luat-ket-hop-thuat-toan-apriori-data-mining/?msclkid=72618846c06d11ec9c0cc395e14265db>

[2] <https://spiderum.com/bai-dang/Khai-pha-Luat-ket-hop-bang-thuat-toan-Apriori-Data-Analytics-for-Business-3-t0rgWnFiVXuj>

[3] <https://viblo.asia/p/khai-pha-mau-pho-bien-va-luat-ket-hop-gGJ59QAa5X2>