

**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO
MÔN THƯƠNG MẠI ĐIỆN TỬ**

Đề tài: Tạo website so sánh giá với web crawler

Giảng viên hướng dẫn: Thầy Văn Đức Sơn Hà
Lớp: IS334.J21

Nhóm sinh viên thực hiện:

1. Bùi Thị Huyền Trân. MSSV: 16521275
2. Nguyễn Thị Kim Yến. MSSV: 16521485
3. Bùi Nguyên Mão. MSSV: 16520724

TPHCM, tháng 5, năm 2019

1 Giới thiệu đề tài:

- Tạo website đơn giản hiển thị kết quả tìm kiếm các sản phẩm ở các trang web: Thế giới di động, Viễn thông A, FPT Shop.
- So sánh giá sản phẩm và đưa ra gợi ý nơi mua hàng với giá tốt nhất.

2 Các sản phẩm tương tự hiện tại và đánh giá thị trường:

- Về thị trường các công cụ so sánh giá, những công cụ phân tích giá đã xuất hiện bắt đầu từ 1995 và đang phát triển đóng góp một phần không nhỏ vào sự phát triển của thương mại nói chung và thương mại điện tử nói riêng. Năm 2005, các dịch vụ so sánh giá ở Anh đã kiếm được 120→ 140 triệu bảng, với mức độ tăng trưởng 30% một năm. Các công ty có sử dụng công cụ so sánh giá dành khoảng 3-5% ngân sách tiếp thị nhưng đạt được nguồn lợi lớn về phát triển số lượng người dùng (dữ liệu UKPRWIRE).
- Ở Việt Nam, vào một vài năm gần đây, thương mại điện tử bắt đầu bùng nổ kéo theo sự xuất hiện của các dịch vụ so sánh giá. các trang web này thường sẽ thống kê dữ liệu từ các nhà bán hàng và đưa ra sắp xếp, gợi ý cho khách hàng mà không trực tiếp bán sản phẩm. Doanh thu sẽ đến từ các thỏa thuận tiếp thị liên kết.
- Hiện nay đã có 1 số trang web có tính năng tương tự phục vụ tìm kiếm giá ở nhiều cửa hàng với nhiều loại sản phẩm. Theo đánh giá chung, các trang web hiện tại đang có những điểm sau:
 - Giao diện người dùng được cập nhật hiện đại dễ sử dụng
 - Giá sản phẩm được cập nhật từ nhiều nguồn với đa dạng các loại mặt hàng, độ chính xác cao
- Tuy nhiên, cũng có những hạn chế:
 - Tuy các trang web có nhiều loại mặt hàng nhưng vẫn còn chia ra các nhóm ngành ví dụ như: nhóm website so sánh hàng tiêu dùng, so sánh giá dịch vụ vận tải, du lịch, dịch vụ bảo hiểm,...
 - Các website hiện tại thường được cập nhật giá dựa trên dữ liệu từ các gian hàng nên dữ liệu giá có thể sẽ được cập nhật chậm hơn và cần có cơ sở dữ liệu để lưu giá sản phẩm trong quá trình hoạt động
 - Các yếu tố về chất lượng sản phẩm hoặc độ uy tín của gian hàng cũng chưa được đảm bảo.

- Thị trường ở Việt Nam hiện tại đang hướng đến chủ yếu các mặt hàng công nghệ, thời trang hoặc một số nhỏ về các sản phẩm dịch vụ vé máy bay, dịch vụ đi lại và một số rất ít về so sánh các dịch vụ bảo hiểm, tín dụng. Nhu cầu người dùng về các công cụ phân tích giá là không hề thiếu và còn rất nhiều ngành hàng chưa có công cụ đánh giá, so sánh như về các dịch vụ giáo dục (so sánh chất lượng, giá các trung tâm tiếng Anh, các cơ sở giáo dục Đại Học,...), y tế (giá, chất lượng các dịch vụ chăm sóc sức khỏe,...),...

3 Nội dung đồ án:

3.1 Công nghệ sử dụng

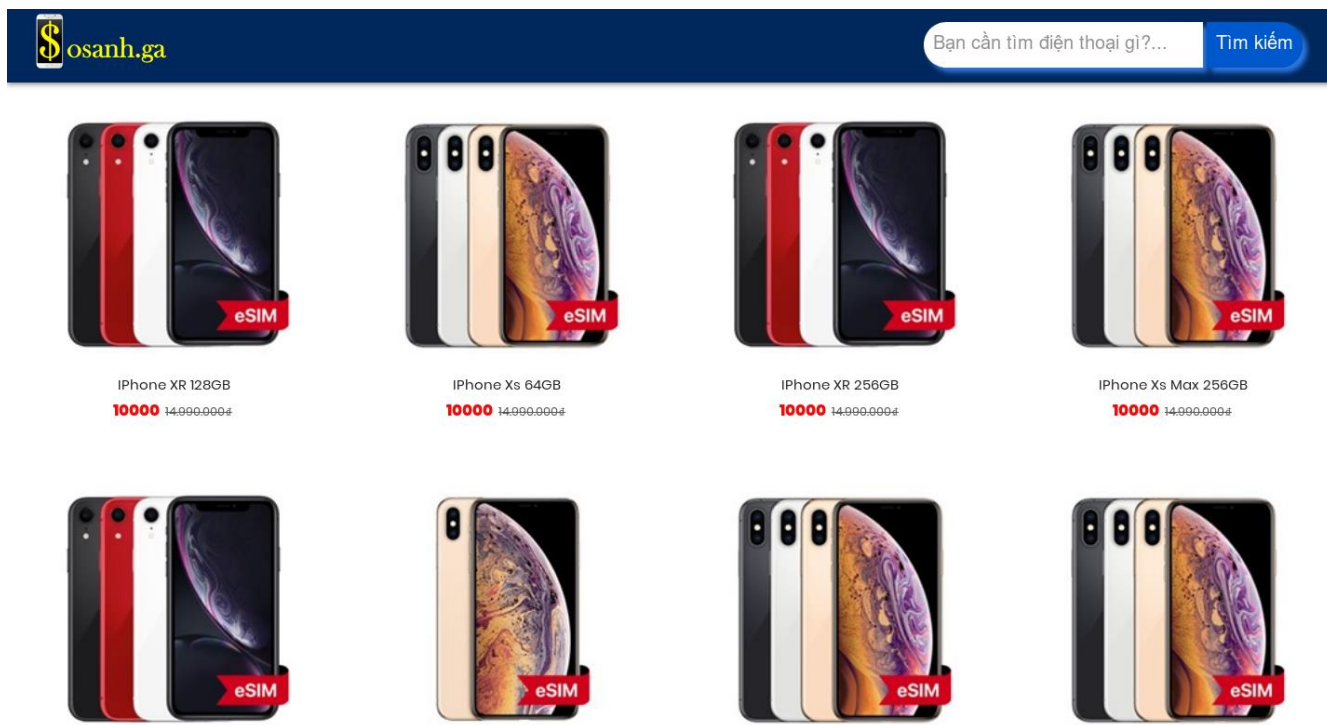
- Xây dựng phần mềm crawler bằng ngôn ngữ python để truy cập website của các gian hàng, phân tích cấu trúc html và lấy thông tin về giá cũng như 1 số đặc tính khác của sản phẩm rồi làm dữ liệu để hiện lên website. Bằng cách này, dữ liệu sản phẩm luôn luôn được cập nhật nhanh, chính xác và không yêu cầu database để lưu trữ. Tuy nhiên sẽ có nhiều khó khăn trong xử lý thông tin thu được từ các trang do có cấu trúc khác nhau và sẽ có ảnh hưởng trong thỏa thuận chiết khấu bán hàng.
- Viết website với python và flask framework để có thể tạo ra một trang web nhẹ, nhanh. Với giao diện đơn giản có mục đích demo các tính năng cơ bản.

3.2 Kết quả đạt được:

- Source code: <https://github.com/nguyenmao2101/sosanh.ga>
- Link trang web: sosanh.ga
- Hoàn thành websites với tính năng tìm kiếm, so sánh, gợi ý các sản phẩm, giao diện hiện đại đơn giản và dễ sử dụng.
 - a) Trang tìm kiếm



b) Trang danh sách sản phẩm



c) Trang so sánh

Bạn cần tìm điện thoại gì?...
 Tìm kiếm

Category 2

New

Bestseller

\$200

\$100

Description

Lorem ipsum dolor sit amet consectetur adipisicing elit. Et dolor suscipit libero eos atque quia ipsa sint voluptatibus! Beatae sit assumenda asperiores iure at maxime atque repellendus maiores quia sapiente.

1

+

-

Add to cart

Vote for your favorite

350 x 250

List group heading

Qui diam libris ei, vidisse incorrupte at mel. His euismod salutandi dissentiunt eu. Habeo offendit ea mea. Nostro blandit sea ea, viris timeam molestiae an has. At nisl platonem eum, Vel et nonumy gubergren, ad has tota facilis probatus. Ea legere legimus tistique cum, sale tantas vim ea, eu vivendo expetendis vim. Voluptua vituperatoribus et mel, ius no elit deserunt mediocrem. Mea facilisi torquatos ad.

14240 votes

Vote Now!

3.3 Kết luận

- Tổng kết lại, thông qua đề án, nhóm đã tìm hiểu được phần nào về thị trường công cụ so sánh giá, công nghệ web crawl và hoàn thành một website so sánh đơn giản quy mô nhỏ.

3.4 Hướng phát triển

- Như đã đánh giá ở trên, thị trường đang có rất lớn và thương mại điện tử Việt Nam đang ngày càng phát triển nên công cụ có rất nhiều hướng để phát triển.
 - Thay đổi ngành hàng hoặc xây dựng với quy mô mặt hàng rộng lớn cả sản phẩm lẫn dịch vụ, từ cả y tế đến giáo dục.
 - Đẩy mạnh các chương trình liên kết với các gian hàng, đảm bảo chất lượng sản phẩm để củng cố lòng tin người dùng và xây dựng thương hiệu như một kênh giới thiệu, đánh giá uy tín, trách nhiệm.
 - Tích hợp thu thập thông tin người dùng, thói quen mua hàng để xây dựng tính năng gợi ý sản phẩm dựa trên công nghệ ML và AI.

4 Nguồn tham khảo:

https://www.ukprwire.com/Detailed/Computers_Internet/Shopping_Comparison_Engines_market_worth_120m-_140m_in_2005_says_E-consultancy_1648.shtml