|  |  |  |
| --- | --- | --- |
|  | BỘ GIÁO DỤC VÀ ĐÀO TẠO  **TRƯỜNG ĐẠI HỌC CÔNG NGHỆ TP. HCM** |  |

**ĐỒ ÁN MÔN HỌC**

**TÌM HIỂU SCRAPY FRAMWORK VÀ SELENIUM, THU THẬP DỮ LIỆU SẢN PHẨM NHÀ THUỐC LONG CHÂU**

Ngành: **KHOA HỌC DỮ LIỆU**

Môn học: **MÃ NGUỒN MỞ TRONG KHOA HỌC DỮ LIỆU**

Giảng viên hướng dẫn : Ths.Lê Nhật Tùng

Sinh viên thực hiện :2286400001 Phạm Quốc An

2286400481 Nguyễn Thị Thu Ngân

TP. Hồ Chí Minh, 2024

|  |  |
| --- | --- |
|  | BỘ GIÁO DỤC VÀ ĐÀO TẠO  **TRƯỜNG ĐẠI HỌC CÔNG NGHỆ TP. HCM** |

**ĐỒ ÁN MÔN HỌC**

**TÌM HIỂU SCRAPY FRAMWORK VÀ SELENIUM, THU THẬP DỮ LIỆU SẢN PHẨM NHÀ THUỐC LONG CHÂU**

Ngành: **KHOA HỌC DỮ LIỆU**

Môn học: **MÃ NGUỒN MỞ TRONG KHOA HỌC DỮ LIỆU**

Giảng viên hướng dẫn : Ths.Lê Nhật Tùng

Sinh viên thực hiện :2286400001 Phạm Quốc An

2286400481 Nguyễn Thị Thu Ngân

TP. Hồ Chí Minh, 2024

**NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN**

TPHCM, Ngày… tháng 10 năm 2024

**Giáo viên hướng dẫn**

(Ký tên, đóng dấu)

**LỜI CAM ĐOAN**

Chúng tôi, Phạm Quốc An và Nguyễn Thị Thu Ngân, xin cam đoan rằng:

Toàn bộ nội dung của bài báo cáo này là kết quả từ quá trình nghiên cứu và làm việc chung của cả hai chúng tôi. Các thông tin được trình bày trong báo cáo đều được thu thập từ các nguồn đáng tin cậy và đã được xử lý cẩn thận.

Chúng tôi đảm bảo rằng không có bất kỳ hành vi sao chép hay sử dụng thông tin không đúng đắn nào từ các nguồn khác. Tất cả các tài liệu tham khảo đã được ghi nguồn rõ ràng và tuân thủ đúng các quy định về trích dẫn học thuật.

Bài báo cáo này là công trình nghiên cứu chung của chúng tôi và chưa từng được nộp hoặc công bố ở bất kỳ đâu trước đây. Chúng tôi cam kết hoàn toàn chịu trách nhiệm về tính trung thực và chính xác của nội dung báo cáo này.

Chúng tôi hy vọng rằng bài báo cáo này sẽ mang đến một góc nhìn tổng quát và chi tiết về chủ đề “...” và đồng thời đóng góp phần nhỏ vào việc nghiên cứu trong lĩnh vực nghiên cứu này.

TPHCM, ngày… tháng 10 năm 2024

**Sinh viên**

Phạm Quốc An,

Nguyễn Thị Thu Ngân

**DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT VÀ TỪ KHÓA**

**MỤC LỤC**

**CHƯƠNG 1: TỔNG QUAN**

**1.1 Giới thiệu đề tài**

Trong thời đại thị trường trực tuyến phát triển mạnh mẽ, việc nghiên cứu và thu thập dữ liệu từ các cửa hàng online đang trở thành một công cụ quan trọng giúp các doanh nghiệp hiểu rõ hơn về cách mà người tiêu dùng hành xử và từ đó tối ưu hóa chiến lược kinh doanh của mình. Đề tài "Phân tích và thu thập dữ liệu trên sàn thương mại Điện Máy Xanh" nhằm mục đích cung cấp một cách tiếp cận khoa học và có hệ thống để hỗ trợ các doanh nghiệp trong việc cải thiện hiệu suất kinh doanh và nâng cao sức cạnh tranh trên thị trường. Chẳng hạn, như việc thu thập dữ liệu về lượng truy cập trang web, thời gian mỗi khách hàng tiêu tại cửa hàng trực tuyến, hoặc thông tin về sản phẩm được mua nhiều nhất. Từ những thông tin này, doanh nghiệp có thể điều chỉnh chiến lược marketing, cải thiện trải nghiệm mua sắm trực tuyến hoặc thậm chí tối ưu hóa dịch vụ khách hàng. Những phân tích và dữ liệu này chính là chìa khóa giúp doanh nghiệp đạt được sự thành công và phát triển bền vững trong thị trường ngày càng cạnh tranh.

**1.2 Nhiệm vụ của đề tài**

Trong dự án "Thu thập và phân tích cửa hàng trực tuyến trên Điện Máy Xanh" chúng ta sẽ sử dụng công nghệ cào dữ liệu để lục lọi thông tin từ trang web của trên Điện Máy Xanh. Điều này giúp chúng ta thu thập dữ liệu về sản phẩm, đánh giá từ khách hàng và các yếu tố khác liên quan. Sau khi thu thập dữ liệu, chúng ta sẽ áp dụng các phương pháp phân tích để khám phá xu hướng tiêu dùng. Ví dụ, chúng ta có thể xem xét cách giá cả ảnh hưởng đến quyết định mua sắm của khách hàng hoặc đánh giá của họ về sản phẩm. Những thông tin này sẽ giúp chúng ta cải thiện chiến lược kinh doanh, tối ưu hóa trải nghiệm mua sắm trực tuyến và tăng cường hiệu quả hoạt động thương mại điện tử.

**1.2.1 Tính cấp thiết của đề tài**

Ngày nay, sự phát triển nhanh chóng của công nghệ và sự phổ biến của thương mại điện tử đã làm thay đổi hoàn toàn cách chúng ta mua sắm. Các cửa hàng trực tuyến không chỉ đơn giản là một nền tảng để mua hàng, mà còn là một khoảng không gian cạnh tranh khốc liệt giữa các thương hiệu. Điện Máy Xanh – một trong những thương hiệu bán lẻ điện tử lớn nhất Việt Nam – đang phải đối mặt với áp lực không nhỏ từ các đối thủ và sự thay đổi liên tục của hành vi người tiêu dùng.

Với quy mô lớn và mạng lưới bán hàng trải rộng trên toàn quốc, Điện Máy Xanh cần tối ưu hóa không chỉ trong chiến lược king doanh mà cvar trong quản lý vận hành, phân phối sản hẩm và cải thiện trải nghiệm khách hàng. Không chỉ có vậy, các yếu tố như giá cả và đánh giá từ người dùng trên các nền tảng thương mại điện tử cũng ngày càng ảnh hưởng mạnh đến quyết định mua hàng. Điều này đặt ra yêu cầu cấp thiết cho Điện Máy Xanh phải tìm ra các phương pháp khoa học và hiệu quả để khai thác dữ liệu, từ đó đưa ra các chiến lược cụ thể nhằm duy trì và phát triển lợi thế cạnh tranh.

Bên cạnh đó, khi thị trường tiêu dùng ngày càng đa dạng và sự kỳ vọng của khách hàng ngày càng cao, việc phân tích dữ liệu lớn từ các hoạt động trực tuyến như đánh giá sản phẩm, hành vi mua sắm trở nên vô cùng quan trọng. Những thông tin này không chỉ giúp doanh nghiệp tối ưu hóa chiến lược marketing mà còn góp phần cải thiện chất lượng dịch vụ, làm hài lòng khách hàng hơn.

Từ những lý do trên, có thể thấy được việc nghiên cứu và triển khai các phương pháp phân tích và thu thập dữ liệu từ các hoạt động thương mại điện tử không chỉ là một yêu cầu cấp thiết, mà còn là chìa khóa để giúp Điện Máy Xanh thích nghi và phát triển trong thời kỳ công nghệ số hóa mạnh mẽ.

**1.2.2 Ý nghĩa khoa học và thực tiễn của đề tài**

Ý nghĩa khoa học: Nghiên cứu này đóng góp vào lĩnh vực khoa học dữ liệu và thương mại điện tử thông qua việc áp dụng các công cụ như web scrapng để phân tích hành vi tiêu dùng. Việc này không chỉ giúp khám phá các xu hướng mới ra ứng dụng tiềm năng trong nghiên cứu marketing và quản trị.

Ý nghĩa thực tiễn: trong thực tiễn, việc phân tích dữ liệu từ Điện Máy Xanh giúp doanh nghiệp tối ưu hóa chiến lược kinh doanh, từ giá cả đến dịch vụ khách hàng. Dữ liệu thu thập được giúp cải thiện trải nghiệm người dùng và nâng cao khả năng cạnh tranh trên thị trường, đồng thời phát hiện và xử lý các vấn đề sớm hơn, giúp tăng doanh thu và hiệu quả vận hành.

**1.3 Mục tiêu**

**1.3.1 Mục tiêu tổng quát**

Đề tài nhằm cung cấp một phương pháp phân tích và thu thập dữ liệu trên sàn thương mại Điện Máy Xanh, giúp hiểu rõ hơn về hành vi tiêu dùng và các yếu tố ảnh hưởng đến quyết định mua hàng. Từ đó đưa ra các đề xuất tối ưu hóa chiến lược kinh doanh, cải thiện trải nghiệm người dùng và nâng cao khả năng cạnh tranh của doanh nghiệp trong thị trường thương mại điện tử đang phát triển mạnh mẽ.

**1.3.2 Mục tiêu cụ thể**

Với bài viết nghiên cứu này, chúng tôi sẽ thu thập dữ liệu về sản phẩm và đánh giá khách hàng trên Điện Máy Xanh, sau đó phân tích hành vi mua sắm để hiểu rõ hơn về xu hướng tiêu dùng và sự ảnh hưởng của các chương trình khuyến mãi. Tiếp theo, nghiên cứu sẽ đánh giá mức độ hài lòng của khách hàng, từ đó đề xuất các chiến lược kinh doanh tối ưu, bao gồm cải thiện dịch vụ, điều chỉnh giá và chiến lược tiếp thi để giúp Điện Máy Xanh nâng cao khả năng cạnh tranh trong thị trường.