**ASSIGNMENT 1 – Phương pháp luận khoa học dữ liệu**

**Chủ đề** : Tăng trưởng số lượng và chất lượng người dùng trên website thương mại điện tử

**Lý do chọn đề tài** :

Bán hàng online trên website thương mại điện tử là xu thế phát triển mạnh mẽ trong thời gian gần đây. Nó ngày càng phát triển, mang lại nhiều lợi ích cho người tiêu dùng, doanh nghiệp và kinh tế nói chung. Dưới đây là 1 số lợi ích khi bán hang online trên website thương mại điện tử :

1. Tiết kiệm thời gian và tiền bạc: Người mua có thể mua hàng hóa và dịch vụ bất cứ lúc nào và ở bất kỳ đâu mà không cần phải di chuyển tới cửa hàng truyền thống. Người bán cũng có thể tiết kiệm được chi phí vận hành cửa hàng.
2. Đa dạng sản phẩm và dịch vụ: website thương mại điện tử cho phép khách hàng có thể truy cập đến hàng nghìn sản phẩm và dịch vụ khác nhau từ nhiều nhà cung cấp trên toàn thế giới.
3. Tăng cường sự tiện lợi: Người tiêu dùng có thể dễ dàng so sánh giá cả và chất lượng của sản phẩm và dịch vụ trên nhiều trang web bán hàng trực tuyến, từ đó có thể đưa ra quyết định mua hàng chính xác hơn.
4. Tăng cường doanh số và khách hàng: website thương mại điện tử cho phép các doanh nghiệp tiếp cận với khách hàng ở khắp nơi trên thế giới, tăng cường doanh số và tạo ra nhiều cơ hội kinh doanh mới.
5. Giảm chi phí vận hành: Doanh nghiệp có thể tiết kiệm được chi phí vận hành, nhân viên và cơ sở vật chất so với mô hình bán hàng truyền thống.
6. Dễ dàng quản lý sản phẩm và dịch vụ: website thương mại điện tử cho phép các doanh nghiệp quản lý sản phẩm và dịch vụ của mình một cách dễ dàng hơn, từ đó có thể cải thiện chất lượng sản phẩm và dịch vụ của mình.

Chính vì có nhiều lợi ích và thế mạnh hơn so với bán hàng truyền thống dẫn đến bán hàng trên website thương mại điện tử hiện đang rất cạnh tranh. Các website thương mại điện tử đều nghiên cứu và phát triển các tích năng tốt nhất để người tiêu dùng ngày càng thoải mái và dễ dàng hơn trong việc mua sắm và trải nghiệm các sản phẩm trực tuyến

|  |  |  |
| --- | --- | --- |
| **Tiêu chí** | **Đặc tả** | **Nội dung** |
| Hiểu biết về Business | Giải thích vấn đề  Miêu tả và xác định các mục tiêu kinh doanh cần đạt được  Khi thực hiện giai đoạn Business Understanding trong một dự án khoa học dữ liệu, cần đặt ra và làm rõ các câu hỏi sau:  Xác định mục đích của dự án  Đặt câu hỏi để hiểu về mục đích: “Làm sao tăng được trải nghiệm mua sắm và kích thích người dùng mua sắm nhiều hơn ?”  Làm rõ vấn đề cần giải quyết | Các mục đích cần đạt được đối với dữ liệu người dùng để tang trưởng số lượng người dùng và doanh thu của website thương mại điện tử :   * Làm sao hiểu được mong muốn, nhu cầu của người dùng khi truy cập website thương mai điện tử ? * Làm sao đưa ra được các sản phẩm gợi ý phù hợp nhất với mong muốn, sở thích của người dùng ? * Làm sao nâng cao được trải nghiệm của người dùng ? * Làm sao để người dùng sau khi truy cập vào website của chúng ta sẽ không muốn truy cập vào các website thương mại điện tử khác ? |
| Hướng tiếp cận giải quyết (Analytic approach) | Các câu hỏi cần trả lời :   * Đầu vào mô hình là gì ? * Mô hình sẽ xử lý những gì ? * Kết quả trả về của mô hình là gì ?   Chọn một trong những hướng tiếp cận phân tích để giải quyết vấn đề:  Predictive model (mô hình dự đoán)  Descriptive model (mô hình miêu tả)  Classification model (mô hình phân loại)  Hoặc mô hình khác | * Đầu vào của mô hình là dữ liệu khảo sát của khách hàng, dữ liệu lịch sử tra cứu, xem hàng, mua hàng, đánh giá sản phẩm. * Mô hình sẽ xử lý dữ liệu lịch sử liên quan đến khách hàng * Kết quả trả về của mô hình, ví dụ cung cấp các loại thông tin :   + Tỉ lệ phần trăm khách hàng hài lòng với sản phẩm.  + Điểm trung bình của khách hàng về các tính năng cụ thể của sản phẩm.  + Phân bố tuổi của khách hàng hài lòng với sản phẩm.  + Tần suất mua sản phẩm của khách hàng hài lòng so với khách hàng không hài lòng.  + Sự khác biệt giữa cảm nhận của khách hàng đối với sản phẩm của nam và nữ.  + Thời gian tìm kiếm sản phẩm của khách hàng  + Tỷ lệ thời gian đặt mua hàng so với thời gian tìm kiếm, đánh giá sản phẩm của khách hàng  + …….  Phương pháp phân tích được sử dụng là mô hình miêu tả |
| Xác định yêu cầu dữ liệu | Miêu tả các yêu cầu dữ liệu để giải quyết vấn đề và cách thức để thu thập dữ liệu.  Tập trung các ý sau:  Loại dữ liệu, số lượng dữ liệu  Nơi thu thập dữ liệu  Có những nguồn dữ liệu nào | Để xác định nội dung của dữ liệu được sử dụng để giải quyết bài toán, một số câu hỏi được đặt ra như sau:  Đầu vào mô hình là gì?  Các yếu tố của dữ liệu cần là gì?  Yêu cầu dữ liệu:  Đầu vào là dữ liệu :   * Lịch sử của khách hàng bao gồm thông tin tra cứu, thời gian đọc sản phẩm, thời gian đặt mua hàng, tỷ lệ phản hồi sản phầm, tỷ lệ hài long sản phẩm. * Dữ liệu phản hồi của người dùng trên các website thương mại điện tử khác   Các yếu tố cần của dữ liệu : Có đầy đủ thông tin thời gian, thao tác khách hàng xử lý trên website thương mại điện tử |
| Thu thập dữ liệu | Nêu phương pháp thu thập dữ liệu được sử dụng  Phân tích về dữ liệu đã thu thập  Thống kê, mô tả về dữ liệu đã thu thập  Xem xét các vấn đề đang có của dữ liệu đã thu thập và lên danh sách khắc phục trong những giai đoạn tiếp theo | Để thu thập dữ liệu :   * Lấy thông tin dữ liệu lịch sử hoạt động của các khách hàng trên database website thương mại điện tử. * Xây dựng tool lấy thông tin dữ liệu phản hồi của người dùng trên các website thương mại điện tử khác   Thống kê dữ liệu   * Thống kê thời gian người dùng truy cập website, thời gian xem thông tin từng sản phẩm * Thống kê các loại sản phẩm mà người dùng hay tìm kiếm, đọc phản hồi * Thống kê tỷ lệ hài lòng của người dùng sau khi mua hàng, trải nghiệm sản phẩm * Thống kê tỷ lệ hài long của người dùng đối với dịch vụ hỗ trợ của website. |