Customer: Customer\_ID; First\_Name; Last\_Name; Postal\_Code; City; Country; Score

Order: Order\_ID; Customer\_ID; Product\_ID; Order\_Date; Shipping\_Date; Sales; Quantity; Discount; Profit; Unit\_Price

Products: Product\_ID; Product\_Name; Category; Sub\_Category

USA Sales: Order ID; Country; Region; State;Sales

DS project

Từ các bảng dữ liệu customer, order và product, bạn có thể áp dụng machine learning vào nhiều mục đích khác nhau để tối ưu hóa các hoạt động kinh doanh của công ty. Dưới đây là một số gợi ý:

1. Dự đoán hành vi mua hàng của khách hàng: Sử dụng lịch sử đơn hàng của khách hàng để dự đoán các mẫu mua hàng tương lai. Điều này có thể giúp bạn tối ưu hóa quảng cáo, cung cấp các ưu đãi cá nhân hóa, và tăng tỷ lệ chuyển đổi.
2. Phân loại sản phẩm: Dựa trên thông tin về sản phẩm như tên, danh mục, và dưới danh mục, bạn có thể phát triển một mô hình phân loại để gắn nhãn tự động cho các sản phẩm mới. Điều này giúp tự động hóa quy trình quản lý sản phẩm và cải thiện trải nghiệm mua sắm cho khách hàng.
3. Phân tích địa lý: Sử dụng dữ liệu về mã bưu chính, thành phố và quốc gia để phân tích xu hướng mua hàng theo vị trí địa lý. Điều này có thể giúp bạn tối ưu hóa kế hoạch vận chuyển và quảng cáo địa phương.
4. Đánh giá khách hàng: Sử dụng dữ liệu khách hàng và đơn hàng để xây dựng một mô hình đánh giá khách hàng dựa trên hành vi mua hàng, số lần mua, và số tiền đã chi tiêu. Điều này giúp bạn xác định và tập trung vào khách hàng có tiềm năng cao nhất.
5. Dự đoán doanh số bán hàng: Sử dụng dữ liệu đơn hàng lịch sử để dự đoán doanh số bán hàng trong tương lai. Điều này giúp bạn lập kế hoạch nguồn lực và dự trữ sản phẩm một cách hiệu quả.
6. Tư vấn sản phẩm: Dựa trên lịch sử đơn hàng của khách hàng và thông tin về sản phẩm, bạn có thể xây dựng một hệ thống tư vấn sản phẩm tự động, giúp khách hàng tìm thấy các sản phẩm phù hợp với nhu cầu và sở thích của họ.

Nhớ rằng, việc áp dụng machine learning đòi hỏi sự chú ý đến quyền riêng tư và tuân thủ các quy định về bảo vệ dữ liệu cá nhân, đặc biệt là khi xử lý thông tin cá nhân của khách hàng.

Chi tiết

1. Dự đoán hành vi mua hàng của khách hàng

Dự đoán hành vi mua hàng của khách hàng là một ứng dụng phổ biến của machine learning trong lĩnh vực thương mại điện tử và bán lẻ. Dưới đây là một số thông tin chi tiết hơn về gợi ý này:

1. Xử lý dữ liệu: Đầu tiên, bạn cần tiền xử lý dữ liệu để chuẩn hóa và làm sạch các thông tin khách hàng và đơn hàng. Điều này có thể bao gồm việc loại bỏ dữ liệu trùng lặp, điền vào các giá trị bị thiếu, và mã hóa các biến phân loại.
2. Tính toán các đặc trưng: Sau khi có dữ liệu sạch, bạn có thể tạo ra các đặc trưng mới từ lịch sử đơn hàng của khách hàng. Các đặc trưng có thể bao gồm số lần mua, tổng số tiền đã chi tiêu, loại sản phẩm thường mua, thời gian giữa các đơn hàng, và các đặc điểm địa lý.
3. Chia dữ liệu: Dữ liệu sẽ được chia thành bộ dữ liệu huấn luyện và bộ dữ liệu kiểm tra. Bộ dữ liệu huấn luyện sẽ được sử dụng để huấn luyện mô hình, trong khi bộ dữ liệu kiểm tra sẽ được sử dụng để đánh giá hiệu suất của mô hình.
4. Lựa chọn mô hình: Bạn có thể sử dụng các thuật toán machine learning như Random Forest, Gradient Boosting, hoặc Neural Networks để dự đoán hành vi mua hàng của khách hàng. Mô hình có thể được huấn luyện để dự đoán xác suất của các sự kiện như việc mua hàng lại trong tương lai, việc mua sản phẩm mới, hoặc việc mua hàng trong một loại sản phẩm cụ thể.
5. Đánh giá mô hình: Sau khi huấn luyện mô hình, bạn cần đánh giá hiệu suất của nó bằng cách sử dụng các phép đo như độ chính xác, độ phủ, và F1-score trên bộ dữ liệu kiểm tra.
6. Tối ưu hóa chiến lược kinh doanh: Cuối cùng, sau khi có một mô hình đủ tốt, bạn có thể sử dụng nó để tối ưu hóa chiến lược kinh doanh của mình. Điều này có thể bao gồm việc tăng cường quảng cáo cho các sản phẩm mà khách hàng có xu hướng mua, cung cấp các ưu đãi cá nhân hóa dựa trên dự đoán của mô hình, và tối ưu hóa quy trình chuỗi cung ứng để đáp ứng nhu cầu dự đoán của khách hàng.

Nhớ rằng, việc thành công của mô hình phụ thuộc vào chất lượng của dữ liệu và sự chọn lựa mô hình phù hợp, cũng như việc thực hiện các bước kiểm định mô hình và tối ưu hóa chiến lược kinh doanh dựa trên kết quả dự đoán.

Các fearture:

|  |  |  |  |
| --- | --- | --- | --- |
| Fearture | | | |
| stt | Fearture | Name of fearture | Thực hiện |
| 1 | Tổng giá trị đơn hàng trước đó của khách hàng | Total\_Order\_Value | Tính tổng giá trị của tất cả các đơn hàng mà mỗi khách hàng đã thực hiện trong quá khứ |
| 2 | Số lần mua hàng của khách hàng | Total\_Order\_Count | Đếm số lượng đơn hàng mà khách hàng đã thực hiện trong quá khứ |
| 3 | Thời gian kể từ đơn hàng cuối cùng của khách hàng | Time\_Sine\_Last\_Order | Tính thời gian kể từ khi khách hàng đã thực hiện đơn hàng cuối cùng |
| 4 | Giá trị trung bình của mỗi đơn hàng | Average\_Order\_Value | Tính giá trị trung bình của các đơn hàng trong quá khứ |
| 5 | Độ phong phú của sản phẩm đã mua | Product\_Diversity | Đếm số lượng sản phẩm khác nhau mà mỗi khách hàng đã mua trong quá khứ |
| 6 | Loại sản phẩm phổ biến nhất | Most\_Common\_Product\_Category | Xác định loại sản phẩm phổ biến nhất mà mỗi khách hàng mua trong quá khứ |
| 7 | Tổng số lượng sản phẩm đã mua | Total\_Products\_Purchased | Tính tổng số lượng sản phẩm đã mua |
| 8 | Thời gian trung bình giữa các đơn hàng | Average\_Time\_Between\_Orders | Tính trung bình các khoảng thời gian giữa các lần mua hàng |
| 9 | Đánh giá trung bình của sản phẩm đã mua | Average\_Product\_Rating | Tính đánh giá trung bình của các sản phẩm mà mỗi khách hàng đã mua |
| 10 | Loại sản phẩm đã mua gần đây nhất | Recent\_Product\_Category | Xác định loại sản phẩm đã mua gần đây nhất trong lịch sử mua hàng |
| 11 | Đặc điểm địa lý | Geographic | Thành phố của khách hàng |
| 12 | Các chiến lược maketing trước đó | Marketing |  |
| 13 | Điểm số của khách hàng | Score |  |
| Label | | | |
| 1 | Tổng giá trị đơn hàng dự kiến cho mỗi khách hàng | Expected\_Total\_Order\_Value | Đây là giá trị mục tiêu mà mô hình sẽ dự đoán. Label được tạo bằng cách tính toán tổng giá trị dự kiến của đơn hàng tiếp theo của mỗi khách hàng trong một khoảng thời gian nhất định (ví dụ: 1 tháng tiếp theo). Điều này có thể được thực hiện bằng cách sử dụng dữ liệu lịch sử đơn hàng để ước tính giá trị trung bình của đơn hàng của mỗi khách hàng, và sau đó áp dụng dự đoán của mô hình cho giá trị này |