



MARKET BASKET ANALYSIS

THE BREAD BASKET Coffee Shop

MỤC TIÊU

Mục tiêu chính của bài báo cáo này là xây dựng một bộ quy tắc bán chéo. Kết quả đầu ra sẽ được sử dụng dựa trên đánh giá của chuyên gia bộ phận nghiệp vụ về khả năng xây dựng thông điệp Marketing tới khách hàng, có thể triển khai dưới dạng thông báo trên Website, ứng dụng; hoặc gửi email, tin nhắn; hoặc triển khai giới thiệu trực tiếp bởi nhân viên kinh doanh.

Sau khi tìm kiếm được một nhóm các sản phẩm có tính liên kết mạnh thỏa mãn các tính chất nêu trên, đơn vị kinh doanh sẽ thực hiện Marketing sản phẩm tới khách hàng, kết quả kỳ vọng rằng sẽ cải thiện được hiệu quả trong việc thu hút khách hàng thực hiện giao dịch.



QUY TẮC BÁN CHÉO

Từ itemset, việc KH sử dụng đồng thời các sản phẩm, tính năng sẽ được xem xét tại từng tháng. Trong đó, mỗi dữ kiện về một KH sử dụng sản phẩm/tính năng tại mỗi tháng là một bản ghi. Mỗi quy tắc kết hợp sản phẩm sẽ bao gồm danh sách sản phẩm tiền đề và danh sách sản phẩm hệ quả. Các quy tắc được đánh giá dựa trên ba hệ số đo lường bao gồm:

- Support
- Confidence
- Lift

Công thức tính các hệ số cụ thể như sau:



QUY TẮC BÁN CHÉO

Xét một quy tắc $\{X\} \rightarrow \{Y\}$ với X là danh sách sản phẩm tiền đề và Y là danh sách sản phẩm hệ quả:

Support. Thể hiện tần suất xuất hiện của itemset trong dữ liệu quan sát.

Support($\{X\} \rightarrow \{Y\}$) = Tổng số bản ghi chứa X và Y / Tổng số bản ghi

Hệ số Support có giá trị nằm trong đoạn $[0, 1]$. Hệ số càng cao thể hiện số lượng quan sát đối có chứa bộ quy tắc càng lớn, thể hiện mức độ tin cậy của bộ quy tắc càng lớn và ngược lại (nếu hệ số này quá nhỏ, thì sẽ có rất ít bản ghi chứa bộ quy tắc đang xét, và khả năng bộ quy tắc được xuất hiện do ngẫu nhiên càng cao, tính tin cậy thấp).



QUY TẮC BÁN CHÉO

Xét một quy tắc $\{X\} \rightarrow \{Y\}$ với X là danh sách sản phẩm tiền đề và Y là danh sách sản phẩm hệ quả:

Confidence: Thể hiện tần suất xuất hiện của sản phẩm Y trong itemset.

$\text{Confidence}(\{X\} \rightarrow \{Y\}) = \text{Tổng số bản ghi chứa } X \text{ và } Y / \text{Tổng số bản ghi chứa } X$

Hệ số Confidence có giá trị nằm trong đoạn $[0, 1]$, giá trị Confidence càng cao thể hiện mức độ liên kết giữa X và Y càng lớn. Trong đó, nếu giá trị này là 1 thì 100% khi khách hàng sử dụng các sản phẩm X , sẽ sử dụng đồng thời sản phẩm Y .



QUY TẮC BÁN CHÉO

Xét một quy tắc $\{X\} \rightarrow \{Y\}$ với X là danh sách sản phẩm tiền đề và Y là danh sách sản phẩm hệ quả:

Lift. Hệ số này so sánh khả năng mua Y kèm điều kiện đã mua X với khả năng mua Y không kèm điều kiện, tức việc mua X khiến khả năng mua Y thay đổi như thế nào.

$$\text{Lift}(\{X\} \rightarrow \{Y\}) = (\text{Số lượng bản ghi chứa } X \text{ và } Y / \text{Số lượng bản ghi chứa } X) / (\text{Số lượng bản ghi chứa } Y / \text{Tổng số bản ghi})$$

Hệ số Lift có giá trị nằm trong khoảng $[0, +\infty)$

- Giá trị Lift = 1 thể hiện X và Y không có ảnh hưởng đến nhau.
- Giá trị Lift < 1 thể hiện ảnh hưởng tiêu cực của hai nhóm hàng, khi khách hàng mua X thì xu hướng mua Y giảm đi.
- Giá trị Lift > 1 thể hiện ảnh hưởng tích cực của hai nhóm hàng, khi mua X thì xu hướng mua Y tăng lên.



KẾT QUẢ MÔ HÌNH

Kết quả đầu ra của mô hình là một số các quy tắc sản phẩm tiền đề và sản phẩm hệ quả tốt nhất thỏa mãn các tiêu chí:

- Các hệ số support, confidence, lift cao
- Số lượng vừa phải (2 - 3 sản phẩm)
- Các sản phẩm, tính năng này phải được tương đối người dùng sử dụng
- Các sản phẩm có tính liên kết tốt để lên chương trình Marketing tới khách hàng
- Các sản phẩm, tính năng này phải có sự khác biệt nào đó đối với các sản phẩm khác được cung cấp bởi đối thủ cạnh tranh





PHÂN TÍCH KHẢ NĂNG BÁN CHÉO CỦA CÁC SẢN PHẨM

THE BREAD BASKET Coffee Shop

TỔNG QUAN

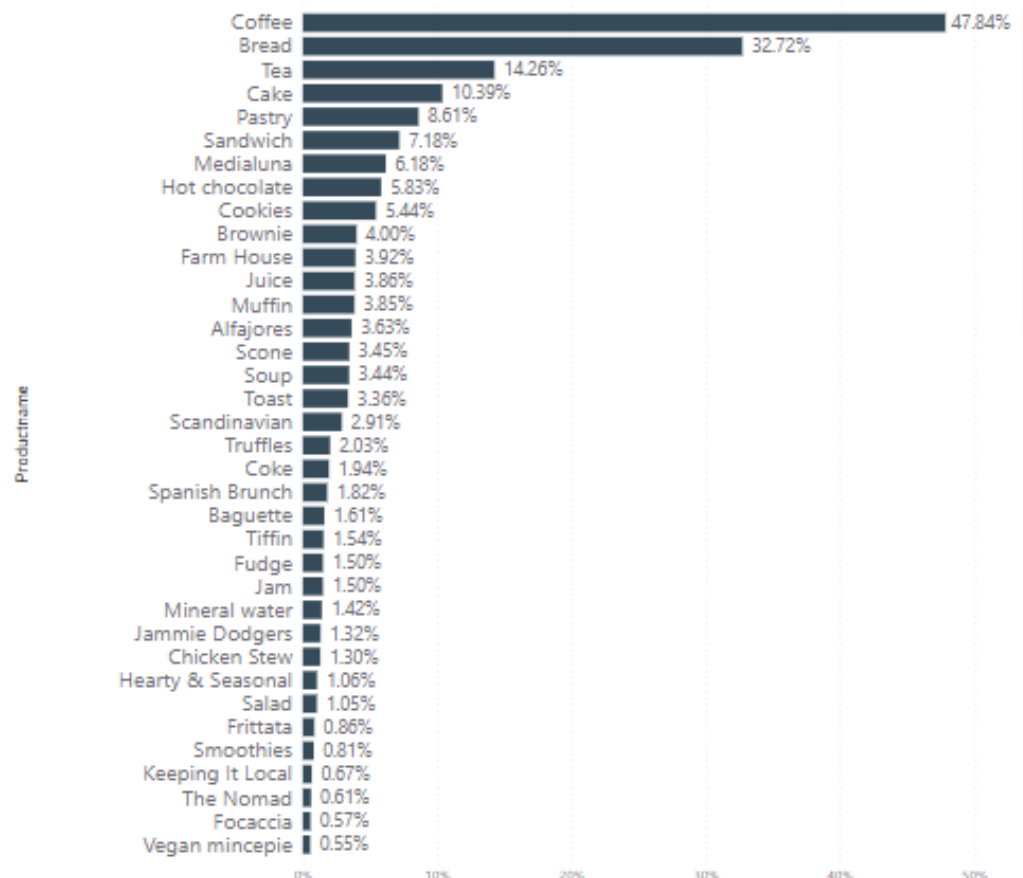
Tổng số sản phẩm của cửa hàng là 94 sản phẩm. Bao gồm:

- Đồ uống như cà phê, trà, sữa,...;
- Các loại bánh mì và bánh ngọt;
- Các loại đồ ăn mặn như sandwich kẹp thịt, trứng, thịt nguội,...

94

Tổng số sản phẩm

Số lượng đơn hàng theo sản phẩm tiến đề



Từ biểu đồ, có thể thấy rằng cà phê và bánh mì là hai sản phẩm bán chạy nhất của cửa hàng, số lượng đơn hàng có cà phê chiếm 47.84% tổng số đơn hàng, và số lượng đơn hàng có bánh mì chiếm 32.72% tổng số đơn hàng.

Đây là hai trong số những sản phẩm rất cơ bản và đặc trưng của một cửa hàng cà phê bánh ngọt. Từ đó ta thấy được cửa hàng đang có những sản phẩm thể mạnh đúng với mục đích kinh doanh.

KHẢ NĂNG BÁN CHÉO

Cà phê là sản phẩm bán chạy nhất của cửa hàng, tuy nhiên nhìn vào bảng dưới, có thể thấy rằng khách hàng ưa chuộng chỉ sử dụng cà phê riêng lẻ mà không mua kèm sản phẩm khác. Vì với cà phê là sản phẩm tiền đề thì xác suất khách hàng sử dụng các sản phẩm đồ ăn và đồ uống khác của cửa hàng là không cao (Confidence dưới 20%) và không có cặp sản phẩm nào quá phổ biến với các khách hàng (Support đều từ 9% trở xuống).

Productname	TSBN chứa A	Productname2	TSBG chứa B	TSBG chứa A và B	TSBG	Support	Confidence	Lift
Coffee	4528	Bread	3097	852	9465	9.00%	18.82%	0.58
Coffee	4528	Cake	983	518	9465	5.47%	11.44%	1.10
Coffee	4528	Tea	1350	472	9465	4.99%	10.42%	0.73
Coffee	4528	Pastry	815	450	9465	4.75%	9.94%	1.15
Coffee	4528	Sandwich	680	362	9465	3.82%	7.99%	1.11
Coffee	4528	Medialuna	585	333	9465	3.52%	7.35%	1.19
Coffee	4528	Hot chocolate	552	280	9465	2.96%	6.18%	1.06
Coffee	4528	Cookies	515	267	9465	2.82%	5.90%	1.08
Coffee	4528	Toast	318	224	9465	2.37%	4.95%	1.47
Coffee	4528	Juice	365	195	9465	2.06%	4.31%	1.12
Coffee	4528	Alfajores	344	186	9465	1.97%	4.11%	1.13
Coffee	4528	Brownie	379	186	9465	1.97%	4.11%	1.03



KHẢ NĂNG BÁN CHÉO

Vì cặp sản phẩm có độ phổ biến cao nhất trong cửa hàng chỉ có 9%, có thể nói rằng cửa hàng hiện tại chưa có cặp sản phẩm nào thật sự được khách hàng ưa chuộng.

Sau khi đặt bộ lọc với hệ số Support từ 2% trở lên, hệ số Confidence từ 40% trở lên, và hệ số Lift lớn hơn 1, ta thu được 4 cặp sản phẩm phổ biến nhất của cửa hàng và có xác suất bán chéo đều từ 50% trở lên:

- Cake + Coffee
- Pastry + Coffee
- Sandwich + Coffee
- Medialuna + Coffee

Productname	TSBN chứa A	Productname2	TSBG chứa B	TSBG chứa A và B	TSBG	Support	Confidence	Lift
Cake	983	Coffee	4528	518	9465	5.47%	52.70%	1.10
Pastry	815	Coffee	4528	450	9465	4.75%	55.21%	1.15
Sandwich	680	Coffee	4528	362	9465	3.82%	53.24%	1.11
Medialuna	585	Coffee	4528	333	9465	3.52%	56.92%	1.19

MỘT SỐ KIẾN NGHỊ ĐỂ TĂNG KHẢ NĂNG BÁN CHÉO

Cả 4 cặp sản phẩm đều là bánh (ngọt hoặc mặn) và sản phẩm hệ quả là cà phê. Có thể nói khách hàng có xu hướng mua bánh và kèm theo một đồ uống, và đồ uống chủ yếu được lựa chọn là cà phê.

- Từ đây, cửa hàng có thể đưa 4 cặp sản phẩm này vào thúc đẩy bán chéo. Mỗi khi có đơn hàng mua một trong các sản phẩm bánh: cake, pastry, sandwich, hoặc medialuna, nhân viên bán hàng có thể gợi ý khách mua kèm cà phê, vì có tới 52 – 57% khách hàng sẽ đồng ý mua thêm.
- Cửa hàng có thể bán 4 cặp sản phẩm này thành các combo có giá ưu đãi hơn mua lẻ, hoặc chạy các chương trình giảm giá cho cà phê khi mua kèm với bánh, hoặc thiết kế các menu ăn sáng/ăn trưa với các cặp sản phẩm này.

Productname	TSBN chứa A	Productname2	TSBG chứa B	TSBG chứa A và B	TSBG	Support	Confidence	Lift
Cake	983	Coffee	4528	518	9465	5.47%	52.70%	1.10
Pastry	815	Coffee	4528	450	9465	4.75%	55.21%	1.15
Sandwich	680	Coffee	4528	362	9465	3.82%	53.24%	1.11
Medialuna	585	Coffee	4528	333	9465	3.52%	56.92%	1.19



DASHBOARD

BÁO CÁO THEO DÕI BÁN CHÉO

Productname
All

Productname2
All

Support: Mức độ phổ biến của cặp sản phẩm A và B.

Confidence: Xác suất khách hàng sẽ mua sản phẩm B sau khi đã mua sản phẩm A.

Lift:

- > 1: Mua sản phẩm A làm tăng khả năng mua sản phẩm B
- < 1: Mua sản phẩm A làm giảm khả năng mua sản phẩm B
- = 1: Cặp sản phẩm A và B không có mối liên hệ

Productname	TSBN chứa A	Productname2	TSBG chứa B	TSBG chứa A và B	TSBG	Support	Confidence	Lift
Bread	3097	Coffee	4528	852	9465	9.00%	27.51%	0.58
Coffee	4528	Bread	3097	852	9465	9.00%	18.82%	0.58
Cake	983	Coffee	4528	518	9465	5.47%	52.70%	1.10
Coffee	4528	Cake	983	518	9465	5.47%	11.44%	1.10
Coffee	4528	Tea	1350	472	9465	4.99%	10.42%	0.73
Tea	1350	Coffee	4528	472	9465	4.99%	34.96%	0.73
Coffee	4528	Pastry	815	450	9465	4.75%	9.94%	1.15
Pastry	815	Coffee	4528	450	9465	4.75%	55.21%	1.15
Coffee	4528	Sandwich	680	362	9465	3.82%	7.99%	1.11
Sandwich	680	Coffee	4528	362	9465	3.82%	53.24%	1.11
Coffee	4528	Medialuna	585	333	9465	3.52%	7.35%	1.19
Medialuna	585	Coffee	4528	333	9465	3.52%	56.92%	1.19
Total	9465		9465		9465			

DASHBOARD

Bộ lọc theo sản phẩm tiền
đề và sản phẩm hệ quả
trong giỏ hàng

Bảng giải nghĩa
các hệ số

Bảng thống kê các sản
phẩm tiền đề, hệ quả,
và các hệ số, có thể
xem kết quả theo các
bộ lọc

BÁO CÁO THEO DỐI BÁN CHÉO								
Productname All		Productname2 All		Support: Mức độ phổ biến của cặp sản phẩm A và B. Confidence: Xác suất khách hàng sẽ mua sản phẩm B sau khi đã mua sản phẩm A. Lift: > 1: Mua sản phẩm A làm tăng khả năng mua sản phẩm B < 1: Mua sản phẩm A làm giảm khả năng mua sản phẩm B = 1: Cặp sản phẩm A và B không có mối liên hệ				
Productname	TSBN chứa A	Productname2	TSBG chứa B	TSBG chứa A và B	TSBG	Support	Confidence	Lift
Bread	3097	Coffee	4528	852	9465	9.00%	27.51%	0.58
Coffee	4528	Bread	3097	852	9465	9.00%	18.82%	0.58
Cake	983	Coffee	4528	518	9465	5.47%	52.70%	1.10
Coffee	4528	Cake	983	518	9465	5.47%	11.44%	1.10
Coffee	4528	Tea	1350	472	9465	4.99%	10.42%	0.73
Tea	1350	Coffee	4528	472	9465	4.99%	34.96%	0.73
Coffee	4528	Pastry	815	450	9465	4.75%	9.94%	1.15
Pastry	815	Coffee	4528	450	9465	4.75%	55.21%	1.15
Coffee	4528	Sandwich	680	362	9465	3.82%	7.99%	1.11
Sandwich	680	Coffee	4528	362	9465	3.82%	53.24%	1.11
Coffee	4528	Medialuna	585	333	9465	3.52%	7.35%	1.19
Medialuna	585	Coffee	4528	333	9465	3.52%	56.92%	1.19
Total	9465		9465		9465			



**THANK YOU
FOR YOUR ATTENTION**