

# XÂY DỰNG MÔ HÌNH THÍ ĐIỂM

- KÍCH HOẠT KHÁCH HÀNG NGỦ ĐÔNG
- BÁN CHÉO/BÁN GIA TĂNG SẢN PHẨM

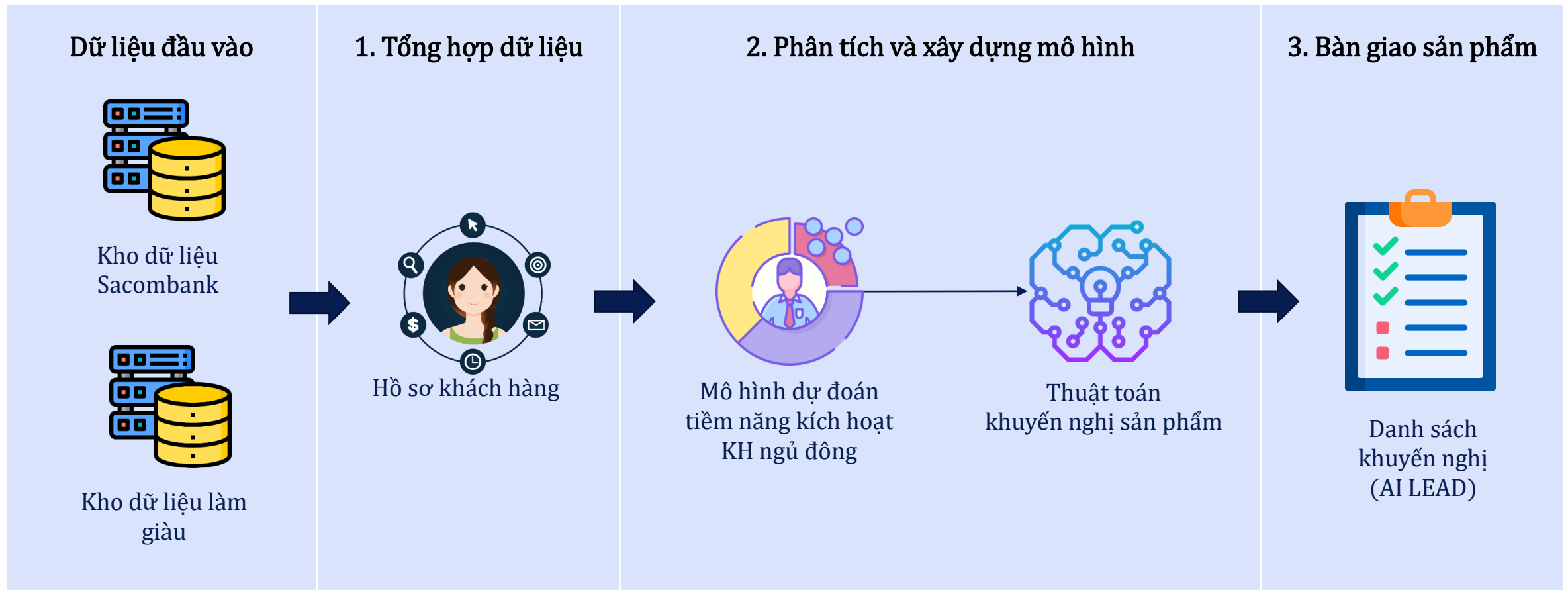
## AGENDA

- 1. PHƯƠNG PHÁP LUẬN
- 2. PHÂN TÍCH FEATURE
- 3. KẾT QUẢ MÔ HÌNH VÀ KẾT QUẢ BACKTEST
- 4. THẢO LUẬN UAT DANH SÁCH LEAD

Sacombank <b>FPT</b> Information System	3
OVERVIEW 1. TỔNG QUAN	
Sacombank <b>FPT</b> Information System	5
DATA PROCESSING 2. XỬ LÝ DỮ LIỆU	
Sacombank <b>FPT</b> Information System	17
MODEL APPROACH 3. ĐỊNH HƯỚNG TIẾP CẬN MÔ HÌNH	
Sacombank <b>FPT</b> Information System	26
TRAINING RESULTS 4. KẾT QUẢ CÁC LẦN CHẠY	
Sacombank <b>FPT</b> Information System	28
IMPLICATIONS OF RESULTS 5. PHÂN TÍCH KẾT QUẢ	
Sacombank <b>FPT</b> Information System	44
NEXT STEP 6. KẾ HOẠCH TIẾP THEO	

# XÂY DỰNG MÔ HÌNH KÍCH HOẠT KHÁCH HÀNG NGỦ ĐÔNG

# 1. PHƯƠNG PHÁP LUẬN



## 2. PHÂN TÍCH FEATURE

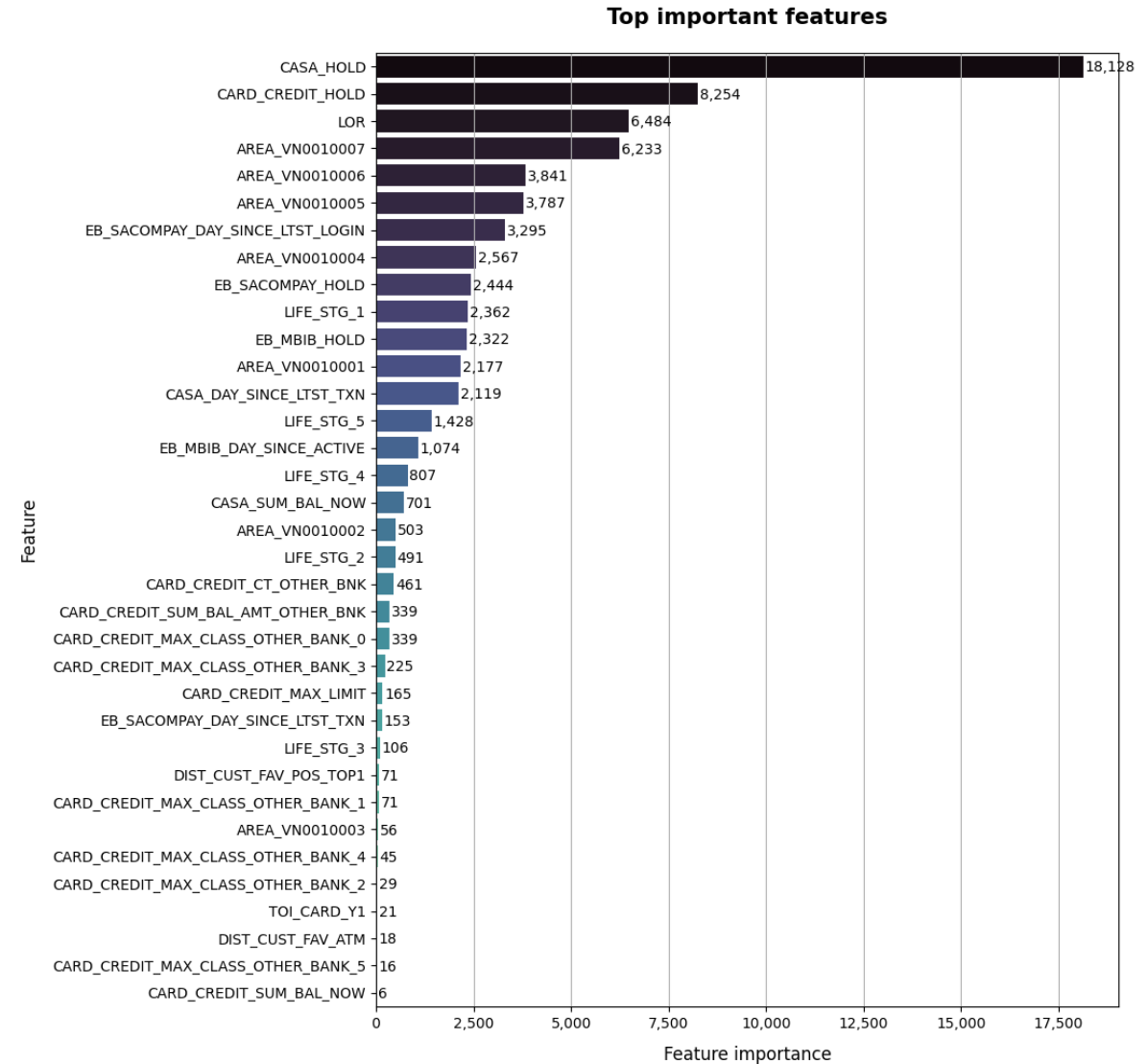
STT	TÊN FEATURE	NHÓM FEATURE	MÔ TẢ
1	LIFE_STG	DEMOGRAPHIC	Giai đoạn cuộc sống
2	LOR	RELATIONSHIP WITH BANK	Thời gian khách hàng gắn bó với ngân hàng
3	AREA	RELATIONSHIP WITH BANK	Khu vực quản lý
4	CREDIT_SCORE	FINANCIAL	Điểm tín dụng
5	CARD_CREDIT_CT_OTHER_BNK	FINANCIAL	Số thẻ tín dụng tại các ngân hàng khác
6	CARD_CREDIT_MAX_CLASS_OTHER_BANK	FINANCIAL	Hạng thẻ tín dụng cao nhất tại ngân hàng khác
7	CARD_CREDIT_SUM_BAL_AMT_OTHER_BNK	FINANCIAL	Tổng số dư khoản tín dụng KH có tại các ngân hàng khác
8	DIST_CUST_FAV_ATM	GEOGRAPHIC	Khoảng cách nơi ở KH tới ATM thường xuyên giao dịch
9	DIST_CUST_FAV_POS_TOP1	GEOGRAPHIC	Khoảng cách nơi ở KH tới top 1 máy POS thường xuyên giao dịch
10	DIST_CUST_FAV_POS_TOP2	GEOGRAPHIC	Khoảng cách nơi ở KH tới top 2 máy POS thường xuyên giao dịch
11	TOI_CARD_Y1	PRODUCT USAGE	Chỉ số TOI thẻ năm ngoái
12	CARD_CREDIT_HOLD	PRODUCT USAGE	1 - Khách hàng sử dụng thẻ tín dụng
13	CARD_CREDIT_INACTIVE	PRODUCT USAGE	Ngủ đông thẻ tín dụng
14	CARD_CREDIT_MAX_LIMIT	PRODUCT USAGE	Hạn mức cao nhất của thẻ tín dụng

### 3. PHÂN TÍCH FEATURE

STT	TÊN FEATURE	NHÓM FEATURE	MÔ TẢ
15	CARD_CREDIT_SUM_BAL_NOW	PRODUCT USAGE	Tổng dư nợ thẻ tín dụng tại thời điểm hiện tại
16	CASA_DAY_SINCE_LTST_TXN	PRODUCT USAGE	Số ngày từ lần giao dịch gần nhất
17	CASA_HOLD	PRODUCT USAGE	1 - KH sử dụng CASA
18	CASA_INACTIVE	PRODUCT USAGE	Ngủ đông CASA
19	CASA_SUM_BAL_NOW	PRODUCT USAGE	Tổng số dư tài khoản thanh toán tại thời điểm hiện tại
20	EB_MBIB_DAY_SINCE_ACTIVE	PRODUCT USAGE	Thời gian từ ngày kích hoạt tài khoản MBIB
21	EB_MBIB_HOLD	PRODUCT USAGE	1 - KH sử dụng MBIB
22	EB_MBIB_INACTIVE	PRODUCT USAGE	Ngủ đông MBIB
23	EB_SACOMPAY_DAY_SINCE_LTST_LOGIN	PRODUCT USAGE	Thời gian từ lần đăng nhập vào hệ thống Sacompay gần nhất
24	EB_SACOMPAY_DAY_SINCE_LTST_TXN	PRODUCT USAGE	Thời gian từ lần thực hiện giao dịch qua hệ thống Sacompay gần nhất
25	EB_SACOMPAY_HOLD	PRODUCT USAGE	1 - KH sử dụng Sacompay
26	EB_SACOMPAY_INACTIVE	PRODUCT USAGE	Ngủ đông Sacompay

## 2. PHÂN TÍCH FEATURE

- Nhóm feature liên quan đến sản phẩm KH sử dụng có độ quan trọng cao nhất.
- Nhóm feature thể hiện mối quan hệ của KH với Ngân hàng (LOR, AREA\_X) cũng có độ quan trọng khá cao.

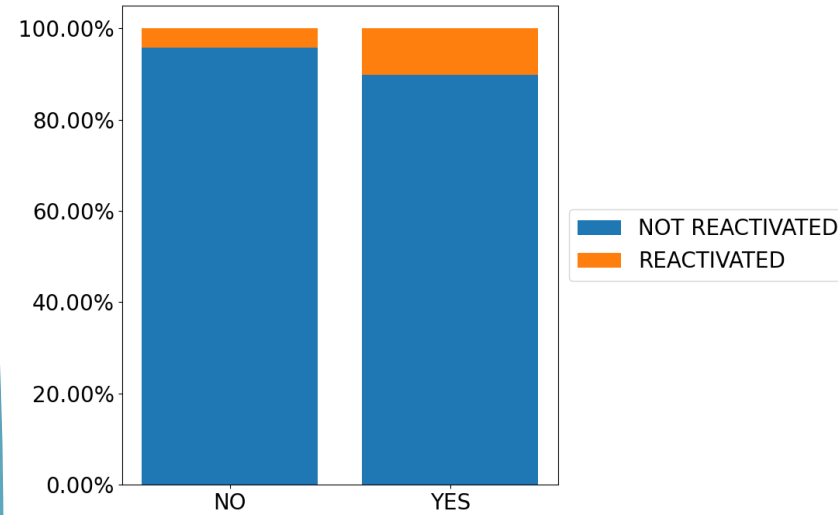




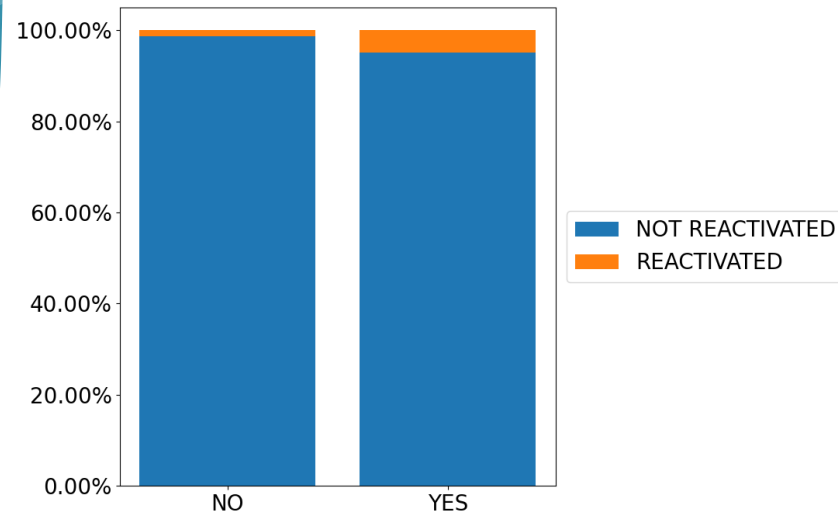
## 2. PHÂN TÍCH FEATURE

- Những KH có thời gian gắn bó với Ngân hàng ngắn hơn có tiềm năng được kích hoạt cao hơn so với những KH đã gắn bó lâu dài.
- Những KH sở hữu sản phẩm CREDIT CARD và CASA có tiềm năng được kích hoạt cao hơn so với những KH không sử dụng hai sản phẩm này.

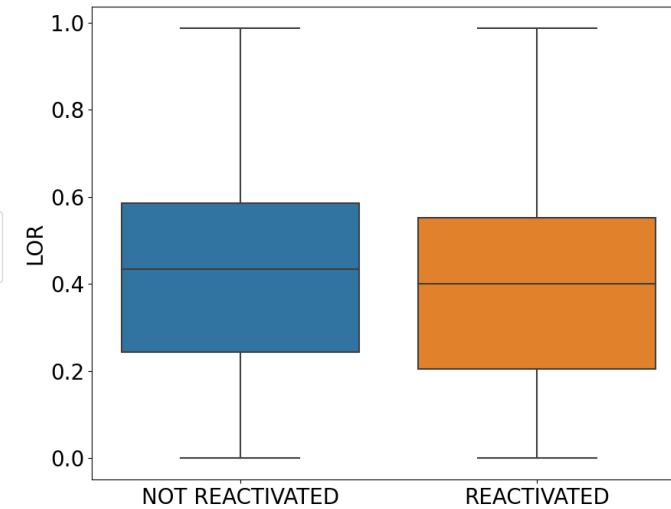
CARD\_CREDIT\_HOLD - Khách hàng sử dụng thẻ tín dụng



CASA\_HOLD - KH sử dụng CASA



LOR - Thời gian khách hàng gắn bó với ngân hàng

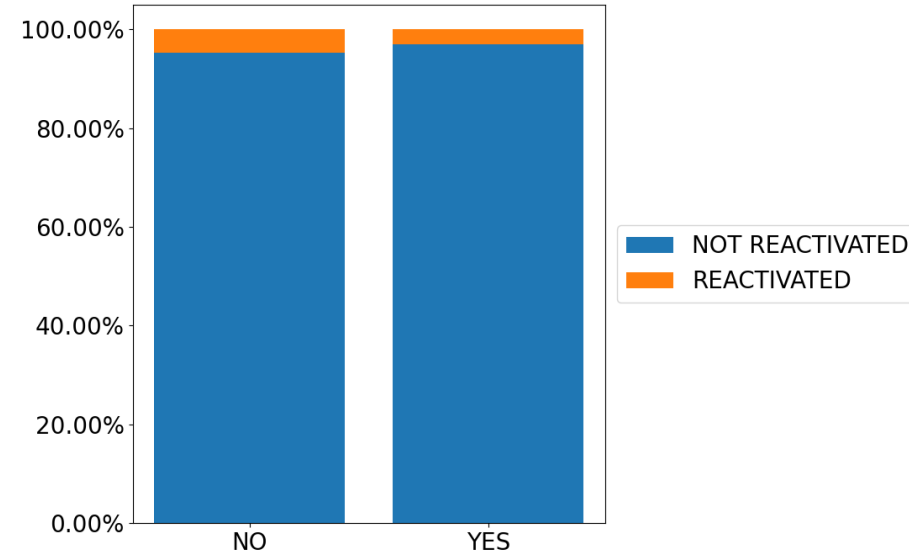




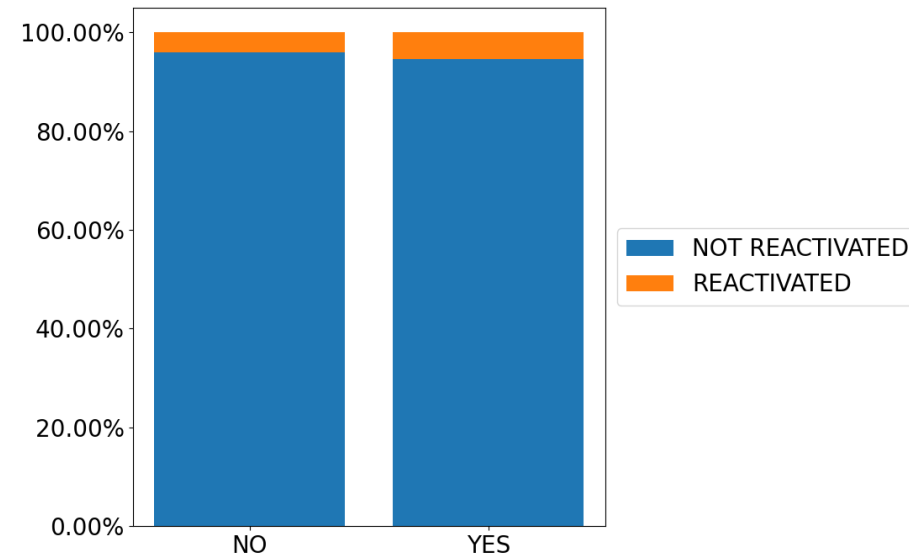
## 2. PHÂN TÍCH FEATURE

- Những KH thuộc KV Tây Nam Bộ có tiềm năng được kích hoạt lại thấp hơn so với các KV khác.
- Trong khi đó, những KH thuộc KV TP.HCM có tiềm năng được kích hoạt lại cao hơn so với các KV khác.

AREA\_VN0010007 - Khu vực quản lý VN0010007



AREA\_VN0010006 - Khu vực quản lý VN0010006

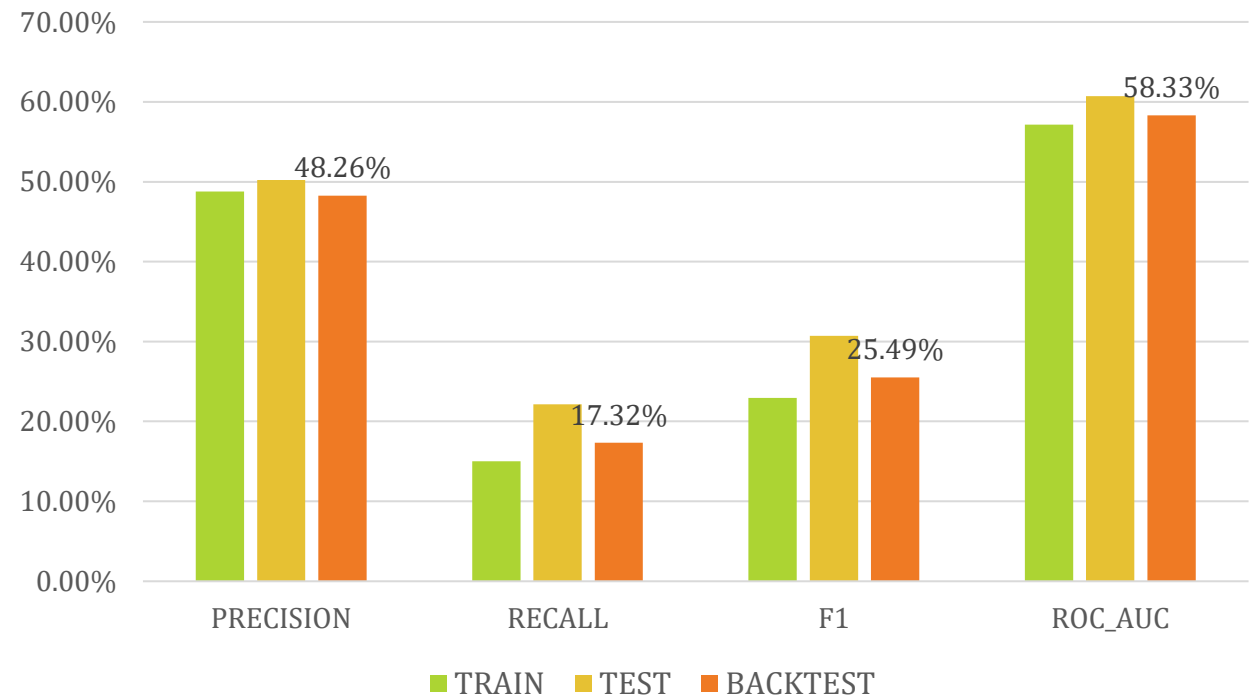


### 3. KẾT QUẢ MÔ HÌNH VÀ KẾT QUẢ BACKTEST

#### PHÂN TÁCH TẬP DỮ LIỆU

TẬP DỮ LIỆU	SLKH	THÁNG					MỤC ĐÍCH SỬ DỤNG
		6	7	8	9	10	
TRAIN	5,548,946						Huấn luyện mô hình
TEST	2,754,627						Kiểm định mô hình
BACKTEST	2,772,006						Backtest
PREDICT	2,794,298						Sử dụng mô hình để dự đoán
Thuật toán: XGBoost Classifier							

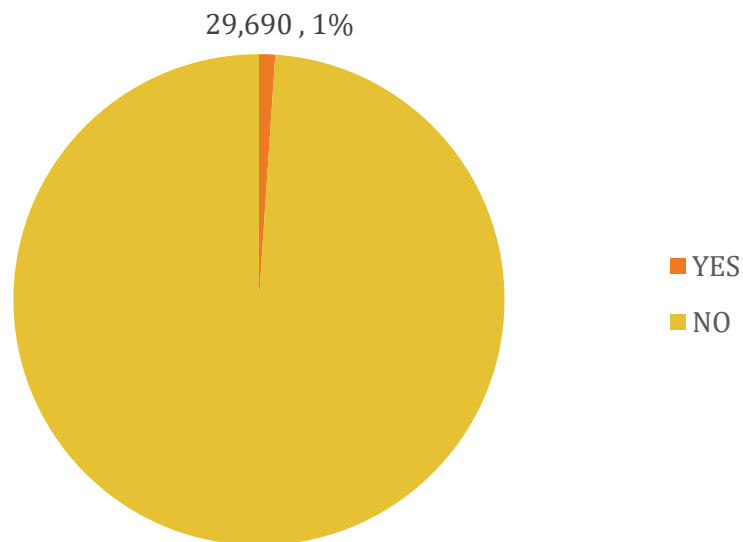
#### ĐÁNH GIÁ MÔ HÌNH



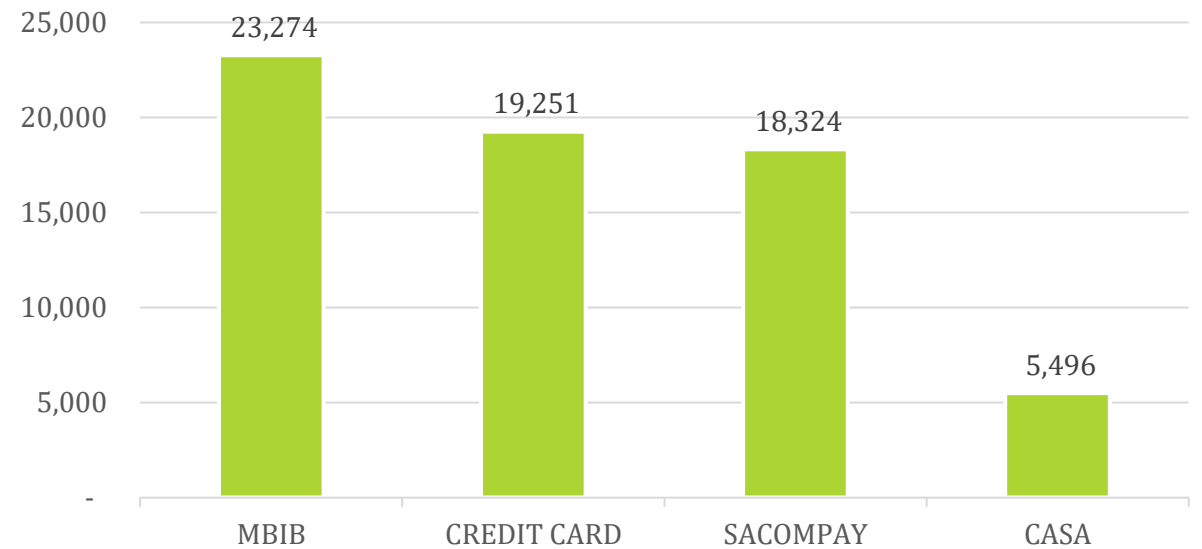
### 3. KẾT QUẢ MÔ HÌNH VÀ KẾT QUẢ BACKTEST

- Kết quả dự đoán tập PREDICT (tháng 10) – danh sách LEAD:

SỐ LƯỢNG KHÁCH HÀNG TIỀM NĂNG



SỐ LƯỢNG KHÁCH HÀNG TIỀM NĂNG THEO SẢN PHẨM



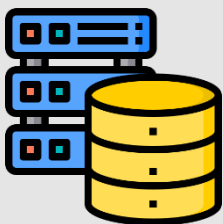
# XÂY DỰNG MÔ HÌNH BÁN CHÉO SẢN PHẨM

*“Định nghĩa một khách hàng được coi là cross-sell:  
Khách hàng mua đầy đủ các sản phẩm trong combo bán  
chéo trong vòng 1 tháng”*

# 1. PHƯƠNG PHÁP LUẬN

14

Dữ liệu đầu vào



Kho dữ liệu  
Sacombank



Kho dữ liệu làm  
giàu

1. Tổng hợp dữ liệu



Hồ sơ khách hàng

2. Phân tích và xây dựng mô hình

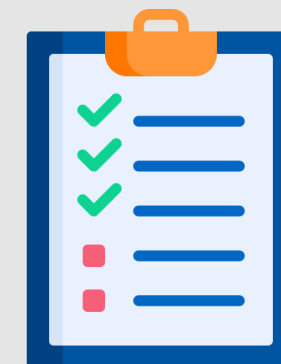


Xác định combo bán chéo  
dựa trên thuật toán  
Apriori



Khuyến nghị sản  
phẩm dựa trên  
combo bán chéo

3. Bàn giao sản phẩm



Danh sách  
khuyến nghị  
(AI LEAD)

## 2. PHÂN TÍCH FEATURE – DANH SÁCH FEATURE ĐẦU VÀO XÁC ĐỊNH COMBO BÁN CHÉO

FTR_NM	GRP	SUB_GRP	DESC
CARD_CREDIT_CROSS_SELL_LABEL1	PRODUCT USAGE	CREDIT_CARD	Thời gian mở thẻ tín dụng thứ nhất
CARD_CREDIT_CROSS_SELL_LABEL2	PRODUCT USAGE	CREDIT_CARD	Thời gian mở thẻ tín dụng thứ hai
CARD_CREDIT_CROSS_SELL_LABEL3	PRODUCT USAGE	CREDIT_CARD	Thời gian mở thẻ tín dụng thứ ba
CARD_CREDIT_CROSS_SELL_LABEL4	PRODUCT USAGE	CREDIT_CARD	Thời gian mở thẻ tín dụng thứ tư
CARD_CREDIT_CROSS_SELL_LABEL5	PRODUCT USAGE	CREDIT_CARD	Thời gian mở thẻ tín dụng thứ năm
CARD_CREDIT_CROSS_SELL_LABEL6	PRODUCT USAGE	CREDIT_CARD	Thời gian mở thẻ tín dụng thứ sáu
CASA_CROSS_SELL_LABEL1	PRODUCT USAGE	CASA	Thời gian mở tài khoản CASA thứ nhất
CASA_CROSS_SELL_LABEL2	PRODUCT USAGE	CASA	Thời gian mở tài khoản CASA thứ hai
CASA_CROSS_SELL_LABEL3	PRODUCT USAGE	CASA	Thời gian mở tài khoản CASA thứ ba
EB_MB_CROSS_SELL_LABEL1	PRODUCT USAGE	MBIB	Thời gian mở tài khoản MB/IB thứ nhất
EB_MB_CROSS_SELL_LABEL2	PRODUCT USAGE	MBIB	Thời gian mở tài khoản MB/IB thứ hai
SACOMBANK_PAY_CROSS_SELL_LABEL1	PRODUCT USAGE	SACOMPAY	Thời gian mở tài khoản Sacompay thứ nhất
SACOMBANK_PAY_CROSS_SELL_LABEL2	PRODUCT USAGE	SACOMPAY	Thời gian mở tài khoản Sacompay thứ hai

Tổng số lượng sản phẩm: 4  
 Tổng số lượng combo cross-sell từ 4 sản phẩm: 11

LOẠI COMBO	SỐ LƯỢNG COMBO	TÊN COMBO
2 sản phẩm	6	(credit, casa) & (credit, MBIB) & (credit, sacompay) & (casa, MBIB) & (casa, sacompay) & (MBIB, sacompay)
3 sản phẩm	4	(credit, casa, MBIB) & (credit, casa, sacompay) & (casa, MBIB, sacompay) & (credit, MBIB, sacompay)
4 sản phẩm	1	(credit, casa, MBIB, sacompay)



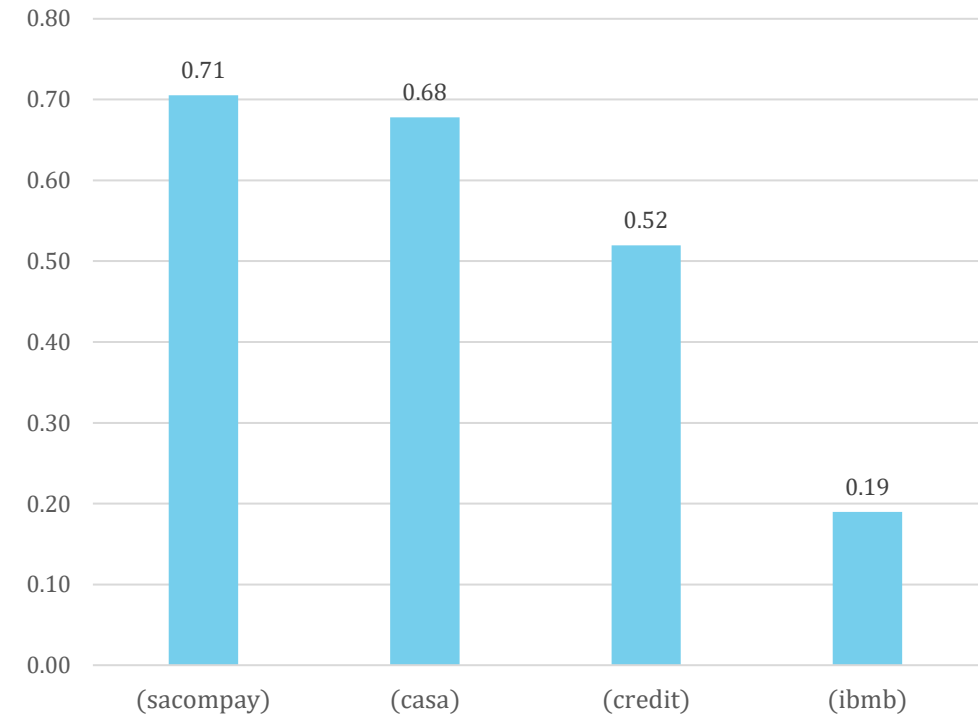
## 2. PHÂN TÍCH FEATURE

Phân tích mức độ phổ biến của 4 sản phẩm.

Mức độ phổ biến được tính trên chỉ số support, được tính như sau:

$$\text{Support}(X) = \frac{\text{số lượng giao dịch sản phẩm } x}{\text{tổng số lượng giao dịch}}$$

Phân bố mức độ phổ biến của 4 sản phẩm



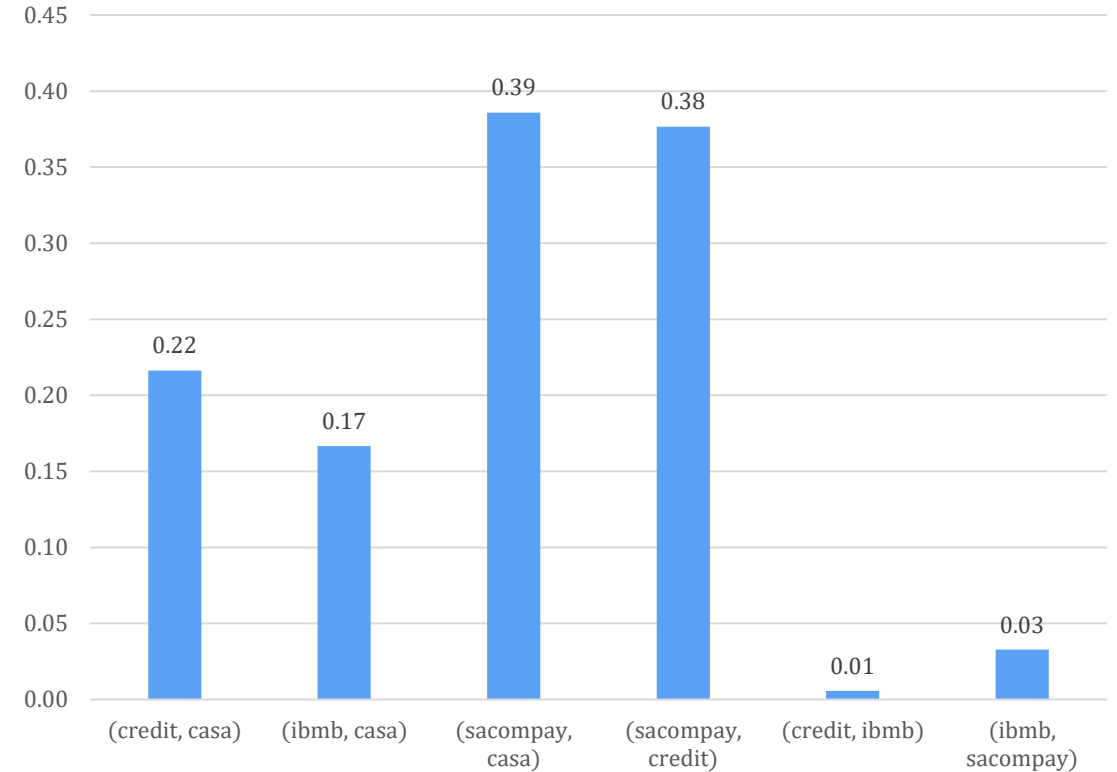
## 2. PHÂN TÍCH FEATURE

Phân tích mức độ phổ biến của combo gồm 2 sản phẩm.

Mức độ phổ biến được tính trên chỉ số support, được tính như sau:

$$\text{Support}(\text{combo } X) = \frac{\text{số lượng giao dịch combo } X}{\text{tổng số lượng giao dịch}}$$

Phân bố mức độ phổ biến của combo 2 sản phẩm



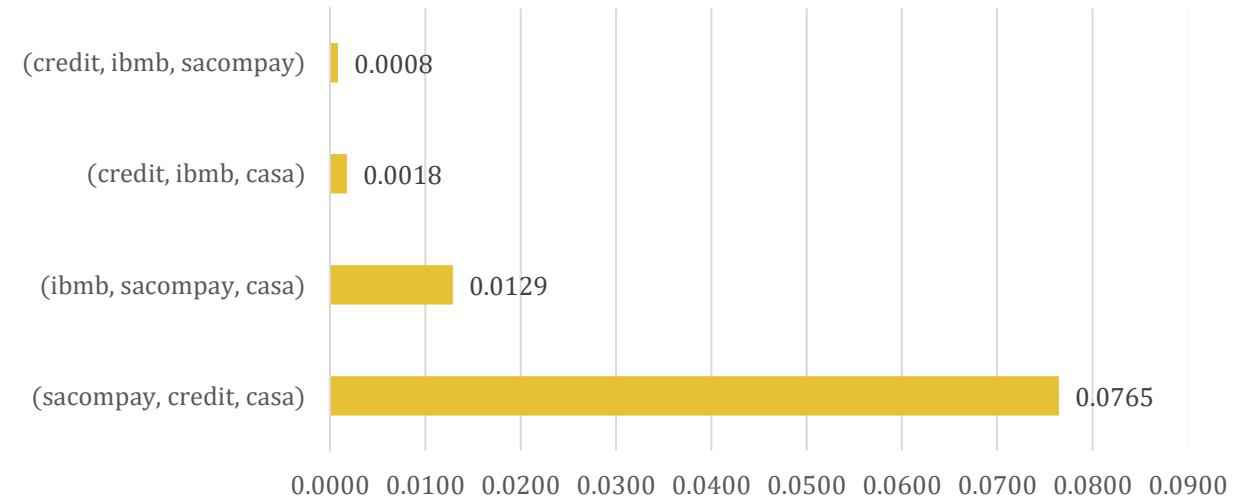
## 2. PHÂN TÍCH FEATURE

Phân tích mức độ phổ biến của combo gồm 3 sản phẩm.

Mức độ phổ biến được tính trên chỉ số support, được tính như sau:

$$\text{Support}(\text{combo } X) = \frac{\text{số lượng giao dịch combo } X}{\text{tổng số lượng giao dịch}}$$

Phân bố mức độ phổ biến của combo 3 sản phẩm





### 3. KẾT QUẢ MÔ HÌNH VÀ KẾT QUẢ BACKTEST

#### KẾT QUẢ MÔ HÌNH

Nhóm Combo	Combo	Support	Confidence	Lift
Hai sản phẩm	(casa, credit) hoặc (credit, casa)	0.22	0.32 hoặc 0.42	0.61
	(ibmb, casa) hoặc (casa, ibmb)	0.17	0.88 hoặc 0.25	1.30
	(sacompay, casa) hoặc (casa, sacompay)	0.39	0.55 hoặc 0.57	0.81
	(sacompay, credit) hoặc (credit, sacompay)	0.38	0.53 hoặc 0.72	1.03
	(credit, ibmb)	0.01	0.01 hoặc 0.03	0.06
	(ibmb, sacompay)	0.03	0.17 hoặc 0.05	0.25
Ba sản phẩm	(sacompay, credit, casa)	0.08	0.20 hoặc 0.35 hoặc 0.20	0.31
	(ibmb, sacompay, casa)	0.01	0.39 hoặc 0.08 hoặc 0.03	0.14
	(credit, ibmb, sacompay)	0.00	0.14 hoặc 0.00 hoặc 0.03	0.01
	(credit, ibmb, casa)	0.00	0.31 hoặc 0.01 hoặc 0.01	0.03

- ▶ Đánh giá kết quả thuật toán Apriori sử dụng 2 chỉ số là Confidence và Lift.
- ▶ Confidence: Xác suất mua sản phẩm Y với điều kiện mua sản phẩm X
- ▶ Lift > 1: Các sản phẩm trong combo combo được mua cùng nhau nhiều hơn kỳ vọng khi được mua riêng lẻ
- ▶ Lift < 1: Các sản phẩm trong combo được ít được mua cùng nhau trong toàn bộ giao dịch

#### COMBO CROSS-SELL:

- ❖ (IBMB, CASA)
- ❖ (CREDIT, SACOMPAY)
- ❖ (CASA, SACOMPAY)

### 3. KẾT QUẢ MÔ HÌNH VÀ KẾT QUẢ BACKTEST

#### KẾT QUẢ BACKTEST

Mô hình	SLKH	SLKH ĐẠT YÊU CẦU	TẬP DỰ BÁO	TẬP THỰC TẾ	DỰ ĐOÁN ĐÚNG TỪ MODEL	DỰ ĐOÁN ĐÚNG TỪ MÔ HÌNH/THỰC TẾ	DỰ ĐOÁN ĐÚNG TỪ MÔ HÌNH/DỰ BÁO
COMBO (IBMB & CASA)	154	154	154	260	153	0.59	0.99
COMBO (CREDIT & SACOMPAY)	2,622	2,622	2,622	708	321	0.45	0.12
COMBO (CASA & SACOMPAY)	64,409	64,409	64,409	15763	14501	0.92	0.23

# XÂY DỰNG MÔ HÌNH BÁN GIA TĂNG SẢN PHẨM

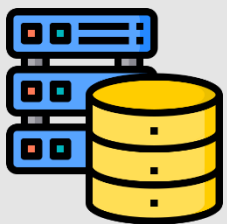


*“Định nghĩa một khách hàng được coi là up-sell: Khách hàng nâng hạng thẻ trong vòng 6 tháng từ thời điểm báo cáo”*

# 1. PHƯƠNG PHÁP LUẬN

24

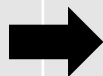
Dữ liệu đầu vào



Kho dữ liệu  
Sacombank



Kho dữ liệu làm  
giàu



1. Tổng hợp dữ liệu



Hồ sơ khách hàng



Lịch sử sử dụng  
sản phẩm của  
khách hàng



2. Phân tích và xây dựng mô  
hình



Mô hình Phân  
loại khách hàng  
Up-sell



3. Bàn giao sản phẩm



Danh sách  
khuyến nghị  
(AI LEAD)

## 2. PHÂN TÍCH FEATURE

STT	TÊN FEATURE	NHÓM FEATURE	MÔ TẢ
1	CARD_CREDIT_AMOUNT_CASH_LESS_30_6M	PRODUCT USAGE	1- KH tỷ trọng doanh số rút tiền mặt không quá 30% tổng doanh số; 0- Còn lại
2	CARD_CREDIT_AMOUNT_SALE_MCC_VANG_50_6M	PRODUCT USAGE	1-KH có doanh số tại MCC vàng chiếm tối đa 50% tổng doanh số; 0- Còn lại
3	CARD_CREDIT_CASH_RATIO_30_6M	PRODUCT USAGE	1- Tỷ lệ thanh toán bình quân tối thiểu từ 30%; 0-Còn lại
4	CARD_CREDIT_NON_DEBT_GRP_6M	PRODUCT USAGE	1- Thẻ phát hành ít nhất 6 tháng và không nợ trong 6 tháng gần nhất; 0-Còn lại
5	CARD_CREDIT_OVER_LIMIT_20_70_6M	PRODUCT USAGE	1- KH có tỷ lệ sử dụng vượt hạn mức từ 20-70%; 0-Còn lại
6	CARD_CREDIT_CT_OTHER_BNK	FINANCIAL	Số thẻ tín dụng tại các ngân hàng khác
7	CARD_CREDIT_MAX_LIMIT	PRODUCT USAGE	Hạn mức cao nhất của thẻ tín dụng
8	CARD_CREDIT_MAX_CLASS	PRODUCT USAGE	Hạng thẻ tín dụng cao nhất
9	CARD_CREDIT_LIMIT_TXN_M3	PRODUCT USAGE	Tỷ lệ giá trị giao dịch trên hạng mức 3 tháng gần nhất
10	CARD_CREDIT_LIMIT_TXN_M6	PRODUCT USAGE	Tỷ lệ giá trị giao dịch trên hạng mức 6 tháng gần nhất
11	CREDIT_SCORE	FINANCIAL	Điểm tín dụng
12	CARD_CREDIT_CT_TXN_1M	PRODUCT USAGE	Số lượng giao dịch bằng thẻ tín dụng trong tháng gần nhất

## 2. PHÂN TÍCH FEATURE

STT	TÊN FEATURE	NHÓM FEATURE	MÔ TẢ
13	CARD_CREDIT_CT_TXN_3M	PRODUCT USAGE	Số lượng giao dịch qua credit card trong 3 tháng gần nhất
14	CARD_CREDIT_SUM_REV_SALE_3M	PRODUCT USAGE	Tổng doanh số sale thẻ tín dụng trong quý gần nhất
15	CARD_CREDIT_SUM_REV_SALE_6M	PRODUCT USAGE	Tổng doanh số sale thẻ tín dụng trong nửa năm gần nhất
16	CARD_CREDIT_SUM_TXN_AMT_3M	PRODUCT USAGE	Tổng giá trị giao dịch trong quý gần nhất
17	CARD_CREDIT_SUM_TXN_AMT_6M	PRODUCT USAGE	Tổng giá trị giao dịch trong nửa năm gần nhất
18	CARD_CREDIT_SUM_BAL_AMT_OTHER_BNK	FINANCIAL	Tổng số dư khoản tín dụng KH có tại các ngân hàng khác tại tháng thứ X trước tháng hiện tại
19	CARD_CREDIT_SUM_BAL_NOW	PRODUCT USAGE	Tổng dư nợ thẻ tín dụng tại thời điểm hiện tại
20	CARD_CREDIT_SUM_TXN_AMT_OTHER_BNK	FINANCIAL	Tổng số tiền giao dịch tại tháng thứ x
21	CARD_CREDIT_UP_SELL_LABEL1_6M	LABEL	1- Khách hàng nâng hạng thẻ trong 6 tháng gần nhất
22	CARD_CREDIT_SUM_REV_CASH_3M	PRODUCT USAGE	Tổng doanh số cash thẻ tín dụng trong quý gần nhất
23	CARD_CREDIT_SUM_REV_CASH_6M	PRODUCT USAGE	Tổng doanh số cash thẻ tín dụng trong nửa năm gần nhất

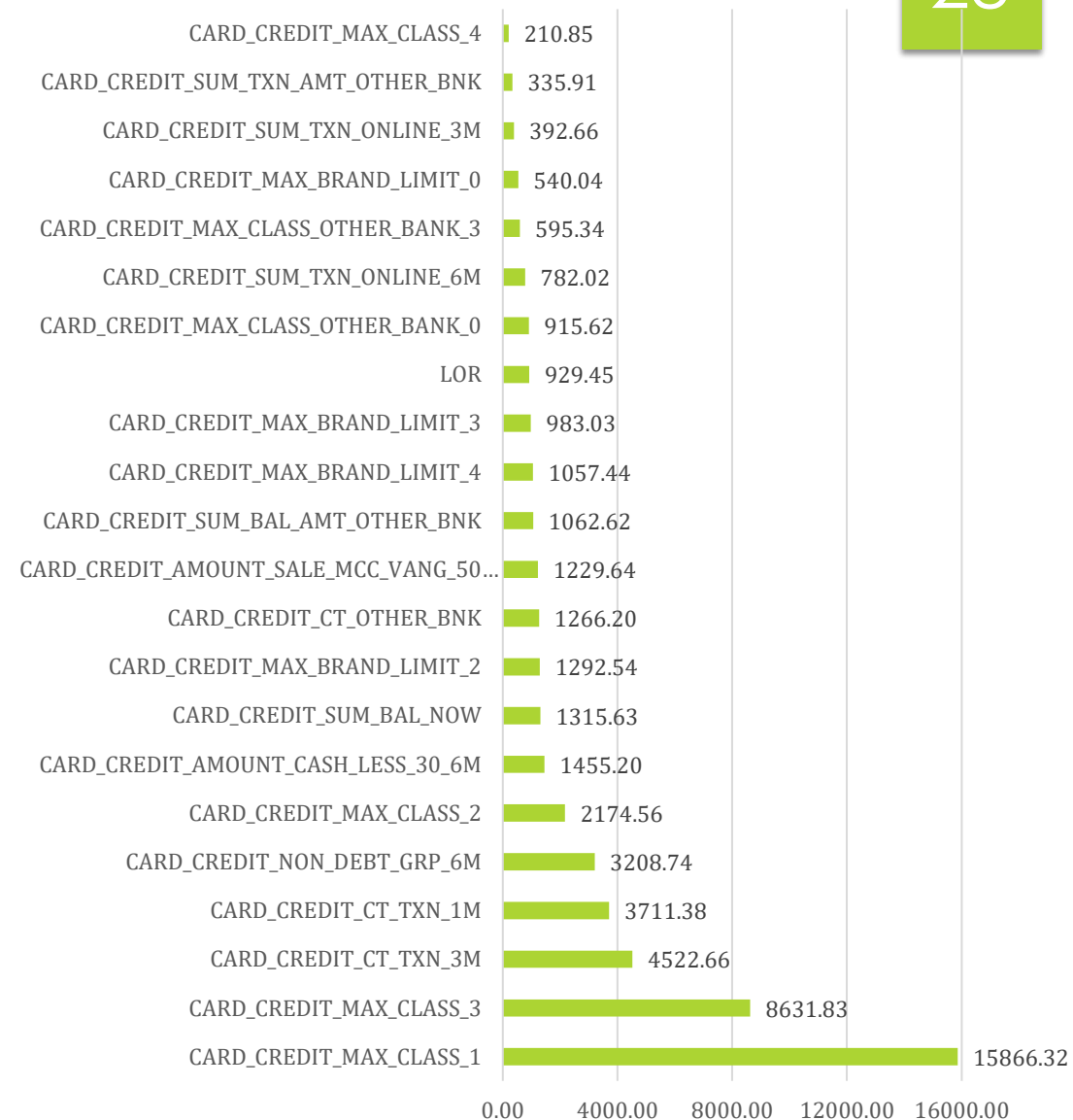
## 2. PHÂN TÍCH FEATURE

STT	TÊN FEATURE	NHÓM FEATURE	MÔ TẢ
24	CARD_CREDIT_SUM_TXN_DOM_3M	PRODUCT USAGE	Tổng giá trị giao dịch Domestic trong 3 tháng gần nhất
25	CARD_CREDIT_SUM_TXN_DOM_6M	PRODUCT USAGE	Tổng giá trị giao dịch Domestic trong 6 tháng gần nhất
26	CARD_CREDIT_SUM_TXN_INTER_3M	PRODUCT USAGE	Tổng giá trị giao dịch International trong 3 tháng gần nhất
27	CARD_CREDIT_SUM_TXN_INTER_6M	PRODUCT USAGE	Tổng giá trị giao dịch International trong 6 tháng gần nhất
28	CARD_CREDIT_SUM_TXN_OFFLINE_3M	PRODUCT USAGE	Tổng giá trị giao dịch offline trong 3 tháng gần nhất
29	CARD_CREDIT_SUM_TXN_OFFLINE_6M	PRODUCT USAGE	Tổng giá trị giao dịch offline trong 6 tháng gần nhất
30	CARD_CREDIT_SUM_TXN_ONLINE_3M	PRODUCT USAGE	Tổng giá trị giao dịch online trong 3 tháng gần nhất
31	CARD_CREDIT_SUM_TXN_ONLINE_6M	PRODUCT USAGE	Tổng giá trị giao dịch online trong 3 tháng gần nhất
32	CARD_CREDIT_MAX_BRAND_LIMIT	PRODUCT USAGE	Thương hiệu thẻ có hạn mức cao nhất
33	CARD_CREDIT_MAX_CLASS_OTHER_BANK	FINANCIAL	Hạng thẻ tín dụng cao nhất tại ngân hàng khác
34	LOR	RELATIONSHIP WITH BANK	Thời gian khách hàng gắn bó vs ngân hàng

## 2. PHÂN TÍCH FEATURE

- Nhóm feature về hạng thẻ hiện tại của khách hàng có tương quan lớn nhất
- Các feature về giao dịch thẻ tín dụng và các feature từ dữ liệu làm giàu có tương quan khá mạnh

Top feature importance

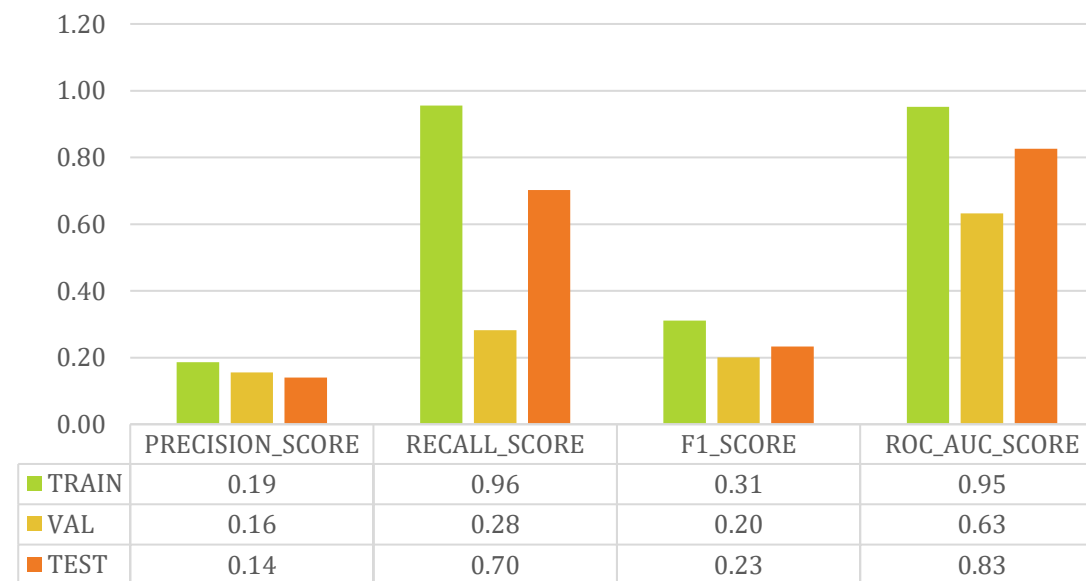


### 3. KẾT QUẢ MÔ HÌNH VÀ KẾT QUẢ BACKTEST

PHÂN CHIA TẬP DỮ LIỆU

TẬP DỮ LIỆU	SLKH	THÁNG					MỤC ĐÍCH
		06	07	08	09	10	
TRAIN	421,750						Huấn luyện mô hình
TEST	442,570						Kiểm định mô hình
BACKTEST	468,912						Backtest
PREDICT	435,623						Dự đoán kết quả
Thuật toán: CatBoost Classifier							

KẾT QUẢ MÔ HÌNH

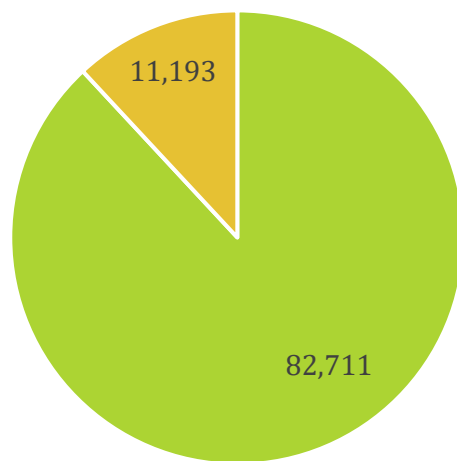




### 3. KẾT QUẢ MÔ HÌNH VÀ KẾT QUẢ BACKTEST

#### KẾT QUẢ TẬP DỰ ĐOÁN – THÁNG 10

Số lượng khách hàng dự báo Upsell



■ No\_Upsell ■ Upsell

#### FORMAT DANH SÁCH LEAD

STT	Mã khách hàng	Tên combo
1	9699738	CASA_SACOMPAY
2	10732253	CASA_SACOMPAY
3	4286084	CASA_SACOMPAY
4	7403288	CASA_SACOMPAY
5	1902770	CASA_SACOMPAY
6	13035851	CASA_SACOMPAY
7	9533788	CASA_SACOMPAY
8	8133091	CASA_SACOMPAY
9	6205340	CASA_SACOMPAY
10	12677431	CASA_SACOMPAY
11	8725677	CASA_SACOMPAY
12	4843358	CASA_SACOMPAY

## 4. THẢO LUẬN UAT DANH SÁCH LEAD