

## Direcionamento de ações de marketing para um produto

Sua empresa lançou um produto há 1 ano e deseja aprimorar suas estratégias de marketing para aumentar a eficácia das campanhas e reduzir os custos envolvidos. O objetivo é identificar o público-alvo mais adequado para as ações de marketing. Com base nisso, você foi convocado para apoiar no processo de análise e tomada de decisão.

### 1. Identificação do Público-Alvo:

Descreva detalhadamente cada etapa do processo que você seguiria para identificar os clientes com maior propensão a adquirir o produto. Explique como você estruturaria a análise e avaliaria a assertividade de suas previsões.

O público-alvo é identificado por uma variável binária (1 para o caso de clientes com maior probabilidade de contratar o produto em questão, 0 para o caso contrário). Consequentemente, trata-se naturalmente de um caso de classificação. O modelo econométrico escolhido para identificar o público-alvo a partir de suas características foi, portanto, o modelo de regressão logística.

Para a aplicação deste modelo, seguiram-se os passos seguintes.

- i. Extração dos dados fornecidos em planilha do Excel e sua organização em DataFrame do pandas, o que resultou na leitura de 2.000 observações.
- ii. Análise descritiva dos dados, por meio de exame de estatísticas descritivas, distribuições (boxplots e histogramas) e correlações.
- iii. Descarte de 678 observações para as quais pelo menos uma coluna apresentava valor faltante, resultando em 1.322 observações.
- iv. Transformação de variáveis categóricas em variáveis *dummy*, de modo a possibilitar a análise de regressão.
- v. Descarte de 5 observações cuja renda mensal constava como negativa.
- vi. Divisão aleatória das 1.317 restantes em um conjunto de dados de treino (987 observações) e um conjunto de dados de teste (330 observações).
- vii. Uso do conjunto de dados de treino para a estimação dos parâmetros de um modelo de regressão logística preliminar que revelou quais das características dos possíveis clientes revelaram-se estatisticamente significantes, a 5%, para a previsão de seu grau de propensão de contratação do produto em questão. O pseudo- $R^2$  deste primeiro modelo foi de 0,7891.

### 2. Análise de Resultados Insatisfatórios:

Suponha que o desempenho do modelo preditivo inicial não tenha sido satisfatório. Detalhe as verificações e diagnósticos que realizaria para identificar possíveis problemas e caso nenhuma falha seja identificada, quais seriam os próximos passos para melhorar os resultados.

Entre os procedimentos seguidos na busca de se elevar o poder preditivo do modelo, incluem-se

- i. Cálculo e testes com variáveis adicionais, tais como o quadrado e o logaritmo das variáveis contínuas originais, na busca de possíveis relações não-lineares entre as variáveis explicativas e a variável explicada.

- ii. Cálculo da variável “número de chamados aberto por ano de relacionamento do cliente com a empresa”, na busca de se mitigar possível colinearidade que poderia reduzir a significância estatística dos parâmetros estimados.
- iii. Eliminação, da modelagem, de variáveis que não se revelaram estatisticamente significantes.

Após a implementação destes passos, observou-se pseudo- $R^2$  de 0,8050, e calculou-se a seguinte matriz de confusão com base no conjunto de dados de teste.

		Predição	
		Positivo	Negativo
Resultado real	Positivo	84	5
	Negativo	11	230

Isto significa que, dos 84 dos 89 clientes potenciais foram corretamente identificados. No caso dos indivíduos que não são clientes potenciais, 230 de 241 deles foram identificados. A precisão do modelo é de  $84 / (84 + 5) = 94,4\%$ , o que significa que 94,4% dos casos positivos foram corretamente identificados. O *recall* do modelo, por outro lado, é de  $84 / (84 + 11) = 88,4\%$ , o que significa que 88,4% dos indivíduos identificados como clientes potenciais são, de fato, clientes potenciais.

### 3. Apoio à Implementação de Ações de Marketing:

Imagine agora que seu modelo alcançou um excelente desempenho na identificação dos clientes mais propensos à contratação do produto. Descreva como você apoiaria o time de marketing na aplicação prática do modelo. Considere desde o planejamento até a apuração de resultados.

Os coeficientes estimados revelam os achados seguintes.

- i. Clientes com rendas mensais maiores são mais propensos a contratar o produto. Ademais, este efeito torna-se ligeiramente mais acentuado em altos níveis de renda.
- ii. Clientes com maior número de acesso ao simulador de preço do produto são mais propensos a adquirirem-no. Quantitativamente, o efeito de um acesso adicional ao simulador na propensão a adquirir o produto equivale ao de aumento de R\$ 158,22 na renda mensal.
- iii. Clientes com maior número de sócios são mais propensos a adquirir o produto. Quantitativamente, o efeito de um sócio adicional equivale ao de aumento de R\$ 4.306,79 na renda mensal.
- iv. Clientes com mais chamados abertos no sistema de suporte tendem a ser menos propensos a adquirir o produto. Quantitativamente, o efeito de um chamado adicional equivale ao de redução de R\$ 8.948,64 na renda mensal.
- v. As demais características dos clientes fornecidas, *i.e.*, idade, gênero, localização geográfica, canal de atendimento mais utilizado, tempo de relacionamento e nota mais recente de satisfação, não revelaram ter efeito estatisticamente significantes na propensão de um indivíduo a adquirir o produto.

Neste contexto, minha recomendação ao time de marketing seria a de priorizar possíveis clientes de acordo com estas características. Adicionalmente, a redução da necessidade de abertura de chamados implicaria, como revela o ponto (iv), expressivo aumento na propensão de clientes médios a adquirirem o produto.

#### 4. Lançamento de um Novo Produto:

Agora, considere que o produto tenha acabado de ser lançado e sua empresa deseja, desde o início, realizar campanhas e ações direcionadas para um público mais específico. Explique o que você mudaria nos três pontos anteriores, descrevendo como você abordaria esse cenário em uma situação onde não há histórico de vendas ou dados relacionados ao produto.

No caso de um produto novo, e portanto da ausência de dados históricos relativos à propensão de clientes a contratá-lo, faz-se necessária nova abordagem.

Uma possibilidade é apoiar-se em dados históricos de um produto já existente, e cujas características sejam o mais similares possíveis às do produto novo. Neste caso, os passos 1-3 descritos acima poderiam ser desenvolvidos com base nos dados históricos do produto já existente e, dada sua similaridade com o novo produto, as conclusões atingidas poderiam, reservadas diferenças entre os dois produtos, ser aplicadas ao público-alvo do novo produto.

#### 5. Variáveis Influentes:

Considere o banco de dados fornecido, contendo as seguintes variáveis:

Propensão\_Contratação: Variável binária indicando se o cliente tem maior probabilidade de contratar o produto (0 - Não, 1 - Sim).

Idade: Idade do cliente.

Gênero: Variável categórica representando o gênero do cliente (Feminino/Masculino).

Região: Localização geográfica do cliente (Norte, Sul, Nordeste, Sudeste, Centro-Oeste).

Qtd\_Acessos\_Simulador: Número de acessos do cliente ao simulador de preço do produto.

Qtd\_Socios: Número de sócios associados a esse cliente.

Renda\_Mensal: Renda líquida média mensal do cliente.

Qtd\_Tickets\_Abertos: Quantidade de chamados abertos por esse cliente no sistema de suporte.

Canal\_Atendimento: Canal mais frequente utilizado pelo cliente para atendimento.

Tempo\_Cliente: Tempo de relacionamento do cliente com a empresa (em anos).

NPS: Nota mais recente de satisfação (NPS) fornecida pelo cliente.

Desenvolva uma análise utilizando Python ou R para identificar as variáveis mais relevantes que contribuem para tornar o cliente mais propenso a contratar o produto.

A análise em Python para identificar as variáveis que contribuem para tornar o cliente mais propenso a contratar o produto encontram-se no repositório

<https://github.com/carlos-e-frohlich/contabilizei>.