

Case Técnico Data Science

Guilherme Nascimento



Proposta do Case:

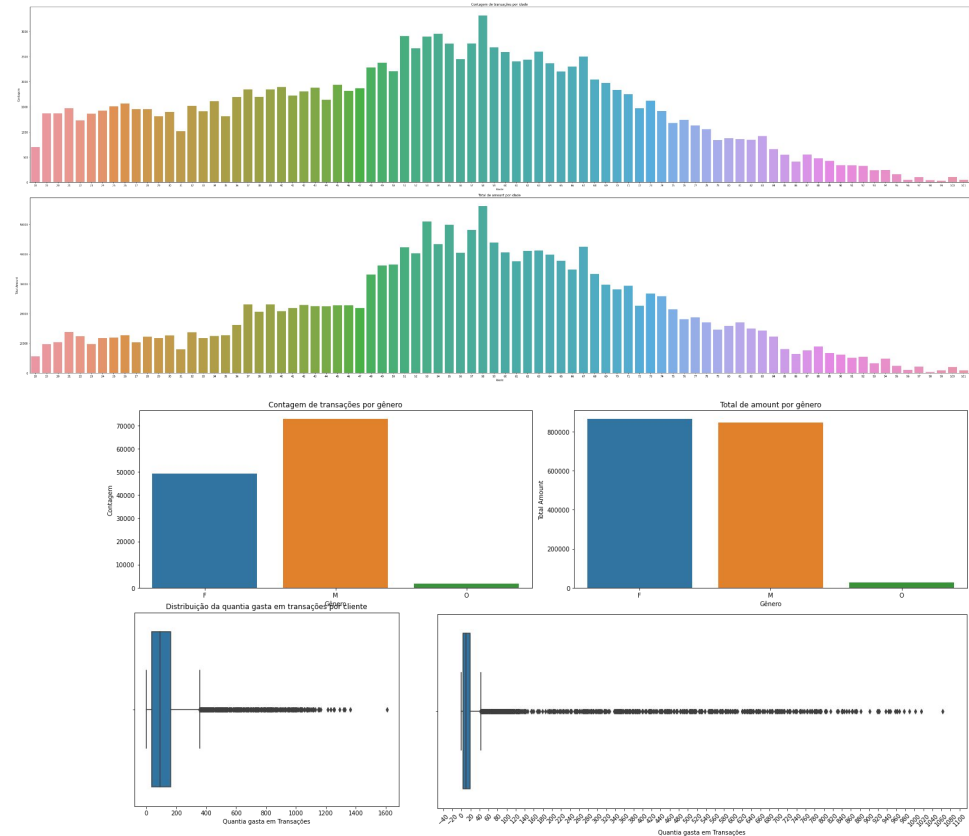
- Desenvolver solução baseada em dados para otimizar a distribuição de cupons e ofertas no iFood.
- O desafio envolve análise de dados, modelagem e comunicação de resultados.
- Cupons e ofertas aumentam engajamento e podem transformar clientes ocasionais em recorrentes.
- Desenvolvimento em Python + PySpark

Resumo da apresentação:

- Análise exploratória dos principais dados considerados na análise
- Modelagem e apresentação dos resultados
- Detalhamento dos possíveis impactos e sugestões de futuros trabalhos e melhorias

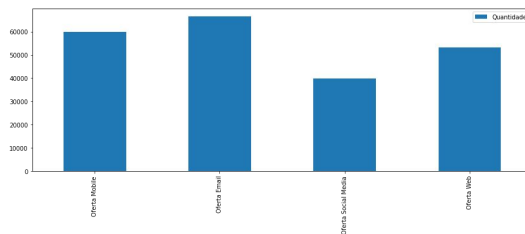
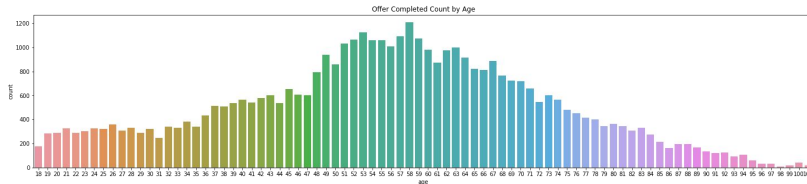
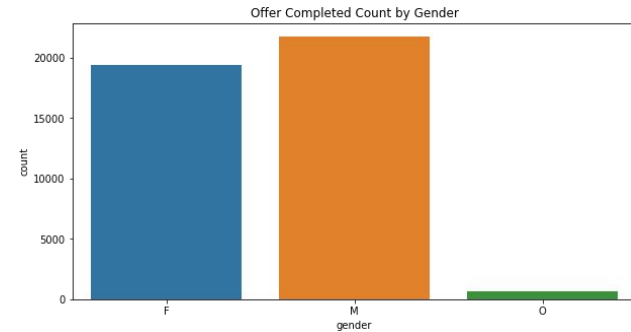
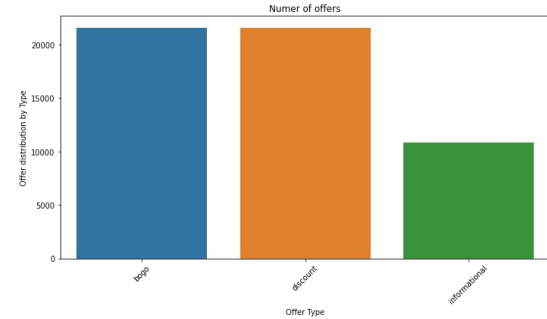
Análise Exploratória: Transações

- Foram feitas análises por idade, gênero e tempo de conta de cada cliente cadastrado na plataforma
- Tanto em quantidade quanto em valor gasto, temos uma maior concentração em clientes de 50-70 anos
- Embora tenha uma diferença no número de transações por gênero, vemos que o total gasto é similar entre homens e mulheres
- A concentração de quantia total gasta por cliente está situado abaixo dos R\$ 200 (quartis 1 e 3) enquanto a concentração dos valores gastos está próximo dos R\$ 20



Análise Exploratória: Ofertas

- Foram considerados gênero, idade, tipo de oferta e veículo de divulgação nas análises
- A efetivação das ofertas segue o mesmo padrão de idade da análise anterior para transações (entre 50-70 anos)
- A efetivação de ofertas entre homens e mulheres tendem a ser similares
- O veículo que mais divulga as ofertas é o email seguido pelo aplicativo
- BOGO (Compre um e leve outro) junto de desconto são as têm um uso similar



Modelagem e resultados

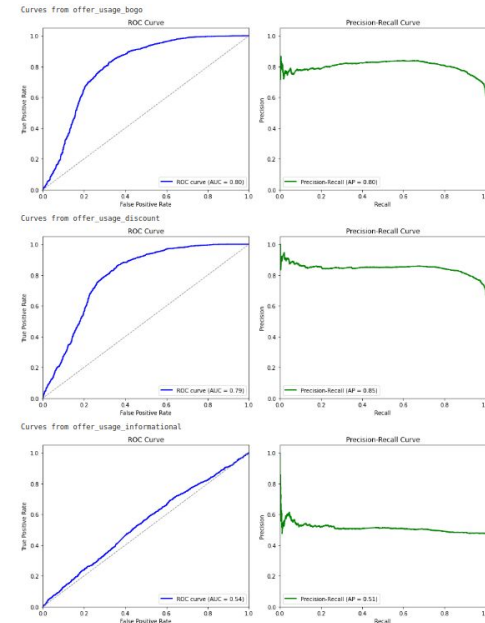
Modelagem

- Nesse case foi considerado como input features as principais informações do cliente (idade, gênero, quantia gasta e suas estatísticas) e as variáveis target as ofertas disponíveis (podendo ser previstas mais de uma para cada cliente)
- As métricas usadas foram F1-score, Precision, Recall e ROC-AUC levando em consideração o acerto para cada categoria.
- Foram considerados 3 algoritmos para avaliação sendo o que teve melhores o SVM.

Resultados

- Bons resultados ao prever as ofertas BOGO e Discount para os clientes porém insatisfatório para ofertas do tipo informational
- Os resultados para informational indicam que o modelo não consegue prever adequadamente, tendo um comportamento aleatório
- Isso mostra que boa parte dos consumidores não tem uma comportamento ou característica definida para interagir com esse tipo de oferta

	Classe	Accuracy	Precision	Recall	F1-score
2	offer_usage_informational	0.533131	0.510804	0.492762	0.501621
0	offer_usage_bogo	0.772777	0.784226	0.866757	0.823428
1	offer_usage_discount	0.790171	0.805038	0.907034	0.852998



Impactos no negócio e melhorias

- A idade do cliente apresenta relevância significativa na análise dos dados de consumo, enquanto o gênero não demonstra diferenças expressivas.
- As ofertas do tipo BOGO e descontos mostram-se eficientes para atrair mais público, porém é importante considerar o valor gasto e a frequência de uso da plataforma para avaliar sua real efetividade.
- As ofertas do tipo *informational* não apresentaram impacto relevante, possivelmente devido à falta de informações que permitam ao modelo fazer previsões precisas. Nesse caso, seria recomendada a reavaliação da estratégia de divulgação ou a inclusão de novos indicadores que possibilitem uma análise mais assertiva dessa oferta.
- Adicionalmente, fornecer dados complementares, como as categorias em que as ofertas ou transações foram utilizadas, poderia enriquecer o modelo, permitindo um direcionamento mais eficiente das campanhas ao público-alvo.

Obrigado!

