



SEGMENTAÇÃO DE CLIENTES DE ALTA RENDA

Entendendo e atendendo clientes com maior poder aquisitivo

Introdução

No mercado financeiro, um dos maiores desafios é avaliar corretamente o perfil de cada cliente para oferecer produtos relevantes e alinhados às suas necessidades.

Visando conectar clientes e produtos, aprimorando o direcionamento das ações de marketing, propomos a adoção de um modelo preditivo de classificação para identificar potenciais clientes de alta renda, baseado em suas características socioeconômicas.



OBJETIVO DO PROJETO

Ferramenta de Machine Learning

Desenvolver ferramenta para identificar clientes com renda anual superior a US\$ 100 mil usando Machine Learning.

Foco em Clientes Qualificados

Concentrar esforços comerciais em prospects com maior potencial de investimento para aumentar conversão.

Otimização de Recursos

Reducir disparos genéricos e otimizar recursos para uma estratégia assertiva orientada por dados.

Relacionamento Personalizado

Fortalecer relações com clientes estratégicos oferecendo soluções personalizadas alinhadas ao perfil econômico.



DESCRIÇÃO E BENEFÍCIOS

Priorização de Clientes

O projeto organiza clientes por probabilidade de alta renda usando dados como vínculo empregatício, escolaridade, faixa etária e histórico financeiro.

Eficiência Comercial

A priorização permite campanhas comerciais mais assertivas, reduzindo custos e aumentando resultados.

Relacionamento Baseado em Dados

Estratégia sofisticada usa insights preditivos para fortalecer a competitividade da instituição financeira.



TRATAMENTO DE DADOS E CONFORMIDADE LEGAL

Processos de Preparação de Dados

Dados foram limpos, padronizados e variáveis agregadas para melhorar a aderência à linguagem de negócios.

Exclusão de Variáveis Sensíveis

Variáveis sensíveis como raça, sexo e país de origem foram removidas para garantir conformidade legal.

Conformidade com LGPD

A conformidade com legislações protege contra discriminação e promove boas práticas internacionais.

Testes de Desempenho do Modelo

Testes internos confirmaram que exclusão de variáveis sensíveis não afetou a performance do modelo.



MODELOS TESTADOS

Diversidade de Algoritmos

Modelos testados incluem versões básicas e versões otimizadas de:

- Regressão Logística
- KNN
- Decision Tree
- Random Forest
- Gradient Boosting
- SVC otimizados

Métricas de Avaliação

Os modelos foram avaliados por acurácia, precisão, recall, F1-score e ROC-AUC, garantindo análise detalhada de desempenho.

Critério de Negócio e Trade-offs

Recall foi priorizado para maximizar captura de clientes ricos, analisando trade-offs com falsos positivos e ROI.

Viabilidade Operacional

Custo computacional e tempo de treinamento foram considerados para garantir implementação eficiente e viável em produção.



RESULTADOS E TRADE-OFFS

Desempenho do Gradient Boosting

Gradient Boosting otimizou o equilíbrio entre F1-Score, ROC-AUC e Recall, sendo ideal para produção.

Precisão do Random Forest

Random Forest apresentou maior precisão, porém com Recall inferior, capturando menos clientes ricos.

Velocidade e Explicabilidade da Decision Tree

Decision Tree é rápida e explicável, ideal para cenários que exigem interpretabilidade do modelo.

Trade-off em Modelos Complexos e Simples

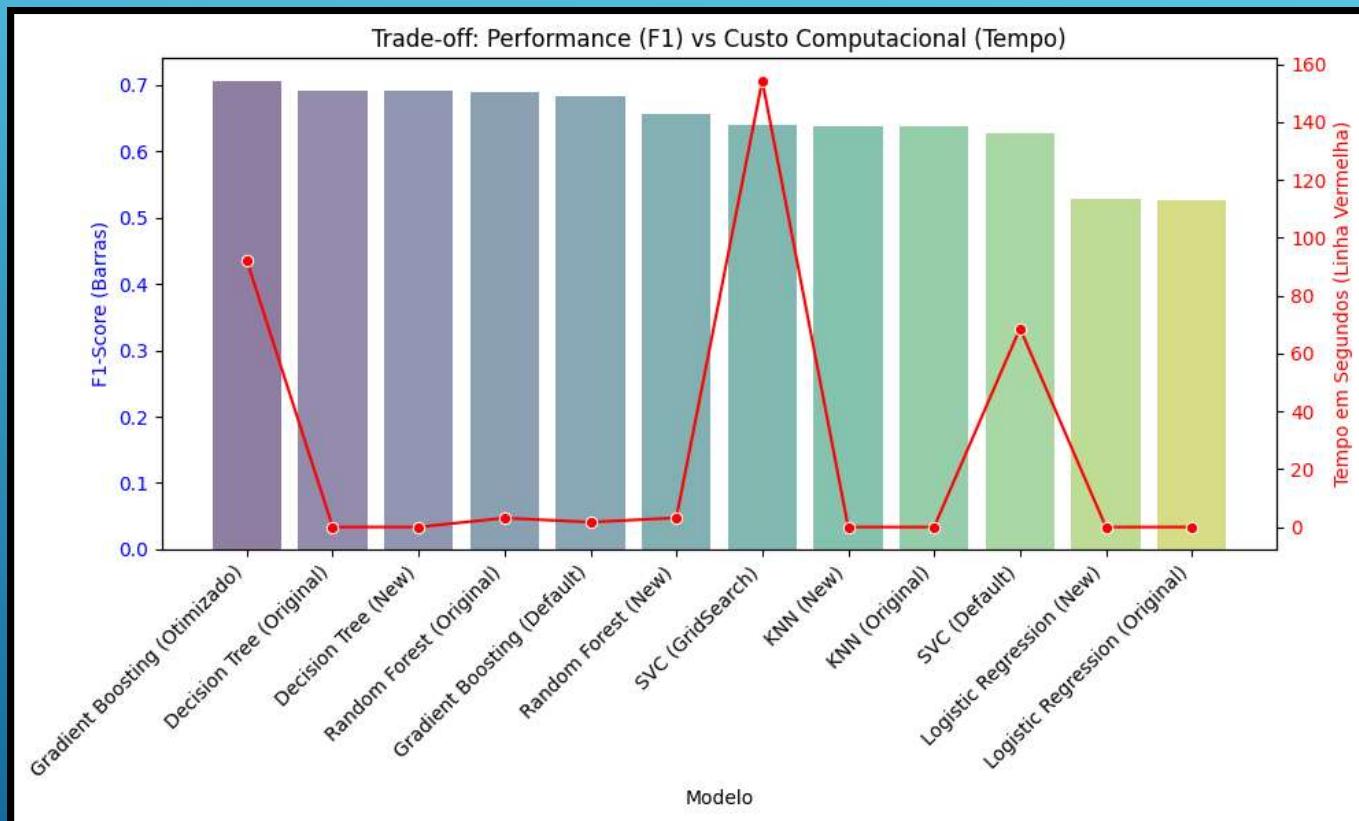
Modelos complexos oferecem maior potencial financeiro, enquanto modelos simples facilitam deploy e explicabilidade.



TABELA DE MÉTRICAS

Modelo	Acurácia	Precisão	Recall	F1-Score	ROC-AUC	Tempo (s)
Gradient Boosting (Otimizado)	0.8633	0.7570	0.6602	0.7053	0.9212	98.34
Decision Tree (Original)	0.8480	0.7088	0.6555	0.6811	0.8984	0.04
Gradient Boosting (Default)	0.8571	0.7704	0.6029	0.6764	0.9124	1.76
Random Forest (Original)	0.8521	0.7972	0.5403	0.6441	0.9161	46.27
KNN (Original)	0.8312	0.6874	0.5837	0.6314	0.8796	45.60
Decision Tree (New)	0.8404	0.7441	0.5420	0.6272	0.8874	0.05
KNN (New)	0.8254	0.6750	0.5687	0.6173	0.8599	22.37
Logistic Regression (Original)	0.8184	0.7279	0.4257	0.5372	0.8499	3.24
Logistic Regression (New)	0.8027	0.6982	0.3578	0.4731	0.8169	0.53

Performance x Custo





DEFINIÇÃO DO MODELO

Modelo Principal Gradient Boosting

Gradient Boosting otimizado é recomendado como modelo principal por seu desempenho superior e potencial de maximizar receita.

Árvore de Decisão como Alternativa

Decision Tree serve como alternativa leve e interpretável, útil para auditorias e suporte explicativo.

Definição de Limiar Comercial

Ajustar limiares de probabilidade para equilibrar recall e ROI, evitando custos com falsos positivos.

Implementação Piloto e Monitoramento

Iniciar com piloto integrado ao CRM para validar hipóteses e ajustar parâmetros com feedback da equipe comercial.



IMPACTO COMERCIAL E CONSIDERAÇÕES

Melhoria no Desempenho Comercial

O Gradient Boosting otimizado identifica 1.191 clientes ricos, aumentando leads qualificados e gerando maior receita com custo operacional controlado.

Mitigação de Riscos e Governança

A estratégia inclui monitoramento contínuo, retreinamento e governança algorítmica com logs, versionamento e auditoria para garantir transparência.

Conclusão e Recomendações

O projeto equilibra desempenho técnico, impacto financeiro e ética, recomendando implantação piloto e evolução para produção orientada a dados.



CONCLUSÃO E BENEFÍCIOS

Ampliação de Clientes de Alta Renda

O modelo amplia a captação de clientes com maior potencial financeiro, aumentando as oportunidades de negócio.

Eficiência nas Ações Comerciais

Equipes comerciais direcionam produtos adequados, reduzindo abordagens irrelevantes e otimizando resultados.

Fortalecimento do Uso de Dados

Uso estratégico de dados melhora o relacionamento com investidores e orienta decisões comerciais.

Inteligência Artificial para Valor

A inteligência artificial atua como ferramenta estratégica para gerar valor e aprimorar a experiência do cliente.