

A person in a white shirt is writing on a notepad with a red pen. In the background, there is a laptop and a tablet. The scene is set on a desk with various office supplies like pens and pencils.

Elo Analytics Case – Detecção de Fraudes para Empresa de Bebidas

Agenda

A stylized, dark-toned illustration of a city skyline at night. It features various skyscrapers, a prominent bridge with multiple cables, and a church with two tall spires. The background is a gradient from dark grey to black.

Nossa
Atuação e
Metodologia

Case
Contexto

Case
Modelos e
Resultados

Case
Próximos
Passos

O desejo de todos atualmente...



“In God we trust.
All others must **bring data**”.
W. Edwards Deming



“Data is the **new oil**”.
Clive Humby

Mas nem todos lembram...

Data Science é ciência: Alguns experimentos dão certo e outros não!

Categoria	Projetos	Em Produção	Resultado (%)
Modelos Probabilísticos	4	1	25%
Modelos Supervisionados	2	1	50%
Modelos Não-Supervisionados	3	1	33%
TOTAL	9	3	33%

Visão dos Papeis num Projeto Completo de Analytics

1



ENQUADRAMENTO

- Entendendo a dinâmica de negócio do cliente, identificar como Analytics pode efetivamente resolver problemas (AI Insight)

2



ENTENDIMENTO & PREPARAÇÃO

- Entender os requisitos, dados e objetivos de negócios
- Coletar e preparar os dados para análise
- Sanity check & EDA preliminar

3



ANÁLISE EXPLORATÓRIA

- Capturar características principais dos dados
- Validar e entender regras de negócios mais profundas
- Balizar estratégias de modelagem

4



MODELAGEM

- Revisar literatura a procura de problemas parecidos
- Explorar diferentes opções de modelagem
- Desenvolver e comparar provas de conceito

5



REFINAMENTO

- Reduzir opções de modelagem para as mais efetivas e explorar mais a fundo
- Ampliar abrangência dos resultados e aumentar rigor científico
- Construir e validar solução de modelagem final

6



ENGENHARIA

- Colocar em produção modelos de alta frequência
- Definir arquitetura da solução e desenvolver integrações e sistemas habilitadores
- Padronizar artefatos de software para conformar com restrições de arquitetura e padrões corporativos

7



MANUTENÇÃO

- Acompanhar resultado e validar contra o esperado
- Realizar ajustes eventuais nos modelos e sistemas
- Incluir novas funções previstas ou suportadas pela arquitetura e corrigir erros eventuais

8



IMPLANTAÇÃO

- Mudar os processos e cultura do cliente para usar a solução desenvolvida em todo seu potencial
- Avaliar a efetividade da solução rodando na prática para um feedback de enquadramento



BA BDA



BA BDA DS



BDA DS



DS



BDA DS



DS DE DEV



DS DE DEV



BA BDA

EQUIPE ENVOLVIDA

Agenda

A stylized, dark-toned illustration of a city skyline at night. It features various architectural elements: a large cathedral with two prominent spires, a suspension bridge with two towers, and several modern skyscrapers of varying heights and shapes. The background is a dark grey, and the buildings are rendered in shades of blue, black, and brown with some yellow highlights for windows or lights.

Nossa
Atuação e
Metodologia

Case
Contexto

Case
Modelos e
Resultados

Próximos
Passos

OBJETIVO DO **MODELO**

APLICAR **TÉCNICAS DE APRENDIZAGEM DE MÁQUINA** PARA **CLASSIFICAR NOTAS FISCAIS DE PRODUTOS PARA PONTOS DE VENDAS COMO FRAUDES.**



OBJETIVO DE **NEGÓCIO**

É ESTIMADO QUE A COMPANHIA **TENHA UMA PERDA DE 10% DO SEU FATURAMENTO** EM FRAUDES. POR ISSO É UM OBJETIVO A AÇÃO DE IDENTIFICAÇÃO DESTES CASOS PARA UMA ATUAÇÃO PROATIVA





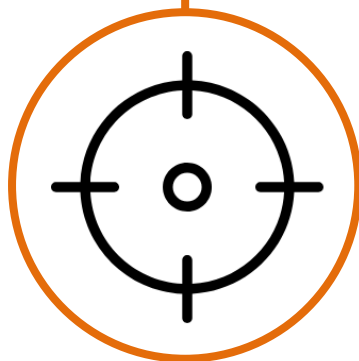
SÍNTESE DO TRABALHO



PROJETO REALIZADO NA **AREA DE ANALYTICS** COM CLIENTE FINAL SENDO A ÁREA DE NEGÓCIOS COM FOCO EM **DECTEÇÃO DE FRAUDES EM VENDAS**



15 MESES DESDE PRIMEIRA CONTRATAÇÃO.
CONTRATAÇÃO INICIAL COMO MANUTENÇÃO DE MODELO EXISTENTE



EVOLUÇÃO DE IMPORTÂNCIA INTERNA COM O CLIENTE.

DUAS RODADAS DE MELHORIA DO MODELO EM PRODUÇÃO

O QUE É UMA FRAUDE?



Representante emite um pedido de vendas para loja A mas venda é na verdade para Loja B



Gera-se uma nota como se pedido fosse para a loja A



Alinhado com o representante o motorista de caminhão entrega os produtos para empresa B



Loja B



Loja A



Loja A

Agenda

A stylized, dark-toned illustration of a city skyline at night. It features various skyscrapers, a prominent bridge with multiple cables, and a church with two tall spires. The background is a gradient from dark grey to black.

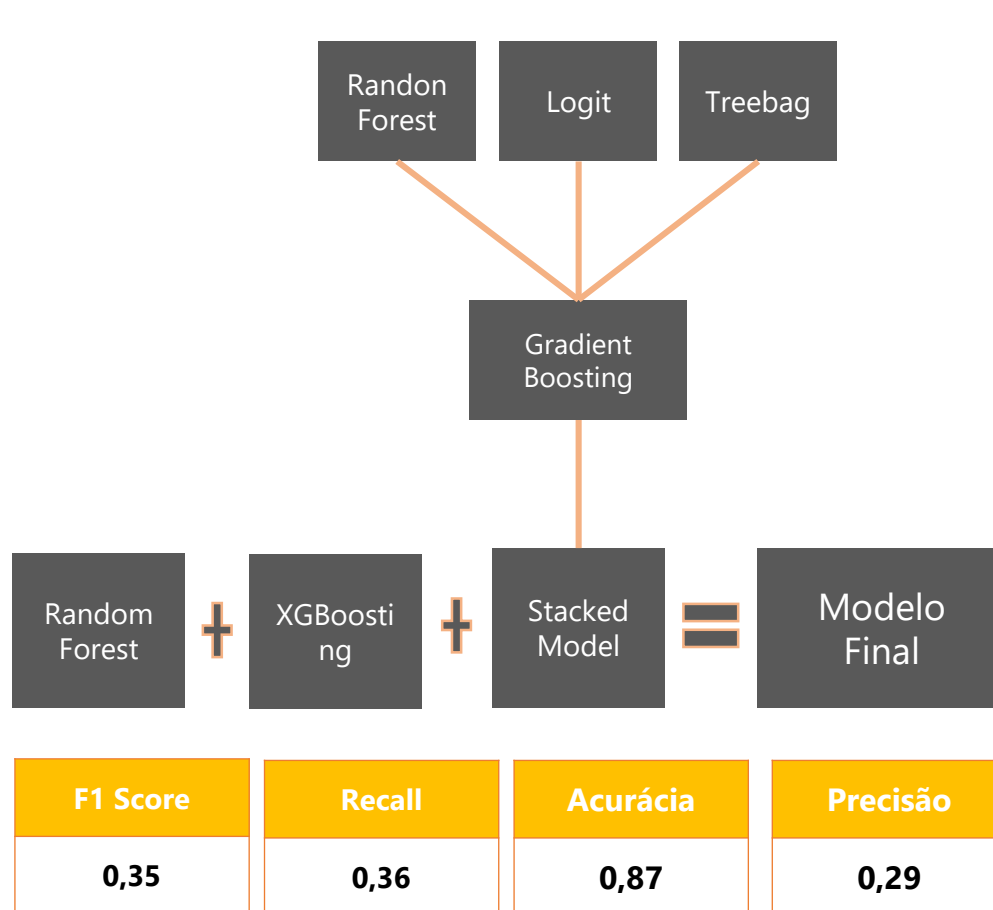
Nossa
Atuação e
Metodologia

Case
Contexto

Case
Modelos e
Resultados

Próximos
Passos

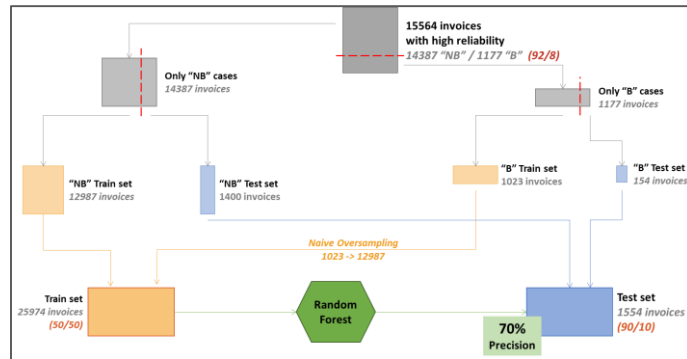
Combinação de Técnicas de Modelagem diferentes



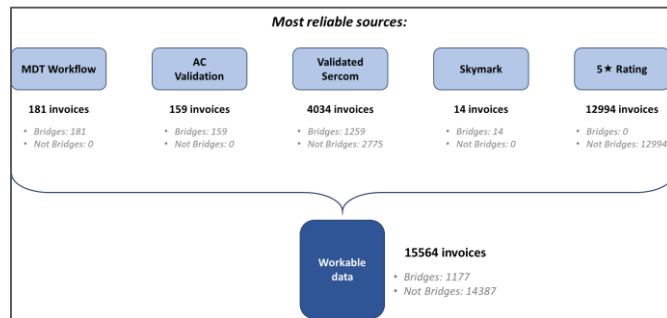
Problemas Identificados

1. **Modelo com alta complexidade e nenhuma interpretabilidade:** não era possível estabelecer um racional de variáveis que impactavam nas fraudes
2. **Objetivo de negócio não era atendido:** métrica de precisão era muito baixa para se ter confiabilidade nas predições feitas
3. Variáveis importantes não eram levadas em consideração

Entendimento e organização do fluxo de apuração de fraudes



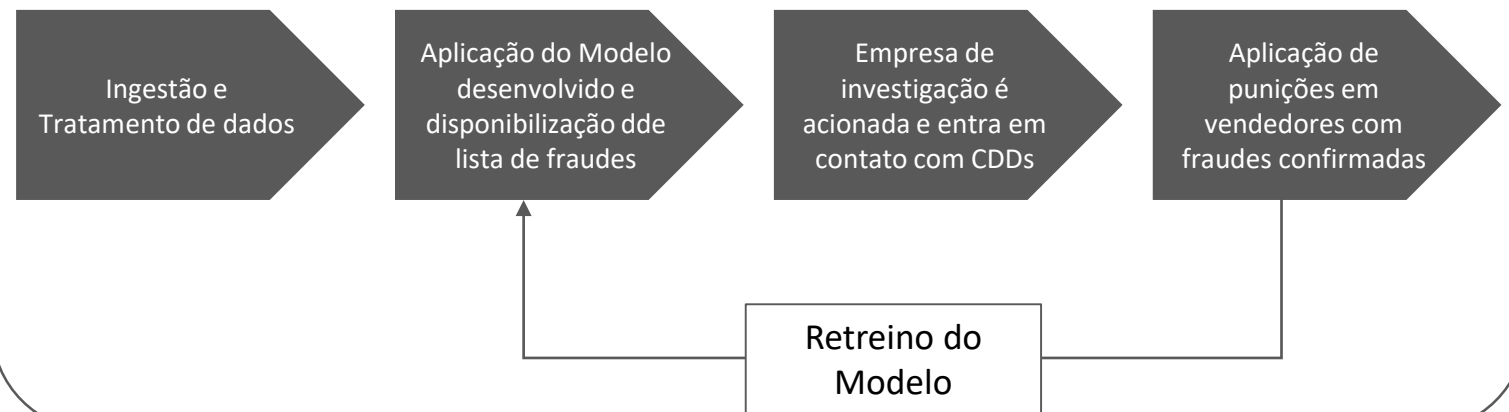
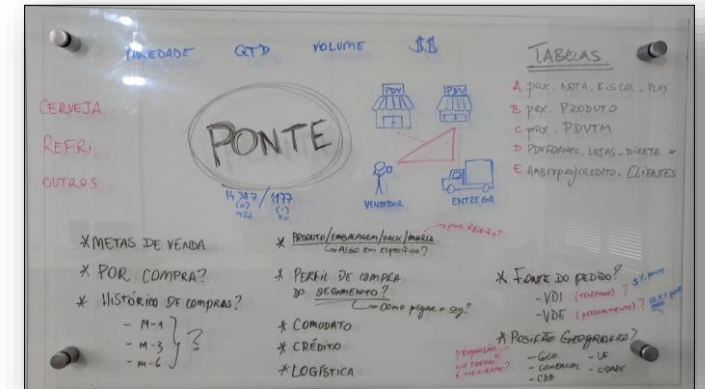
Mapeamento de dados e aplicação de regras de negócio



Criação de uma nova base de treinos a partir de 5 diferentes fontes de classificação

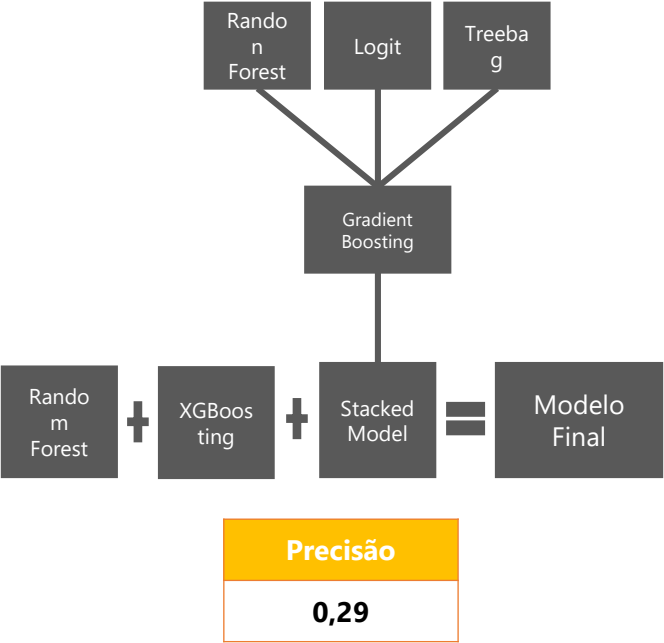
Imersão time de negócios da cervejaria + cientistas de dados Elo Group + equipe de negócios Elo Group

- ✓ *Envolvimento de Vendas e Logística*
- ✓ *Entendimento de todo o processo de vendas*
- ✓ *Análise do perfil das pontes indetificadas*
- ✓ *Experiência de campo*
- ✓ *Entendimento do problema e pesquisa de benchmarkigns*



Evolução dos Modelos de Classificação de Fraudes

Modelo do centro de Analytics India



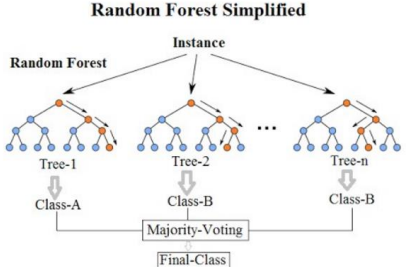
Modelo exploratório



- **Técnica utilizada:** Logit
- ETL para Features via Azure ML
- Utilização de 66 variáveis
- **Precisão:** 60% (Base Teste não Balanceada)

Precisão
0,59

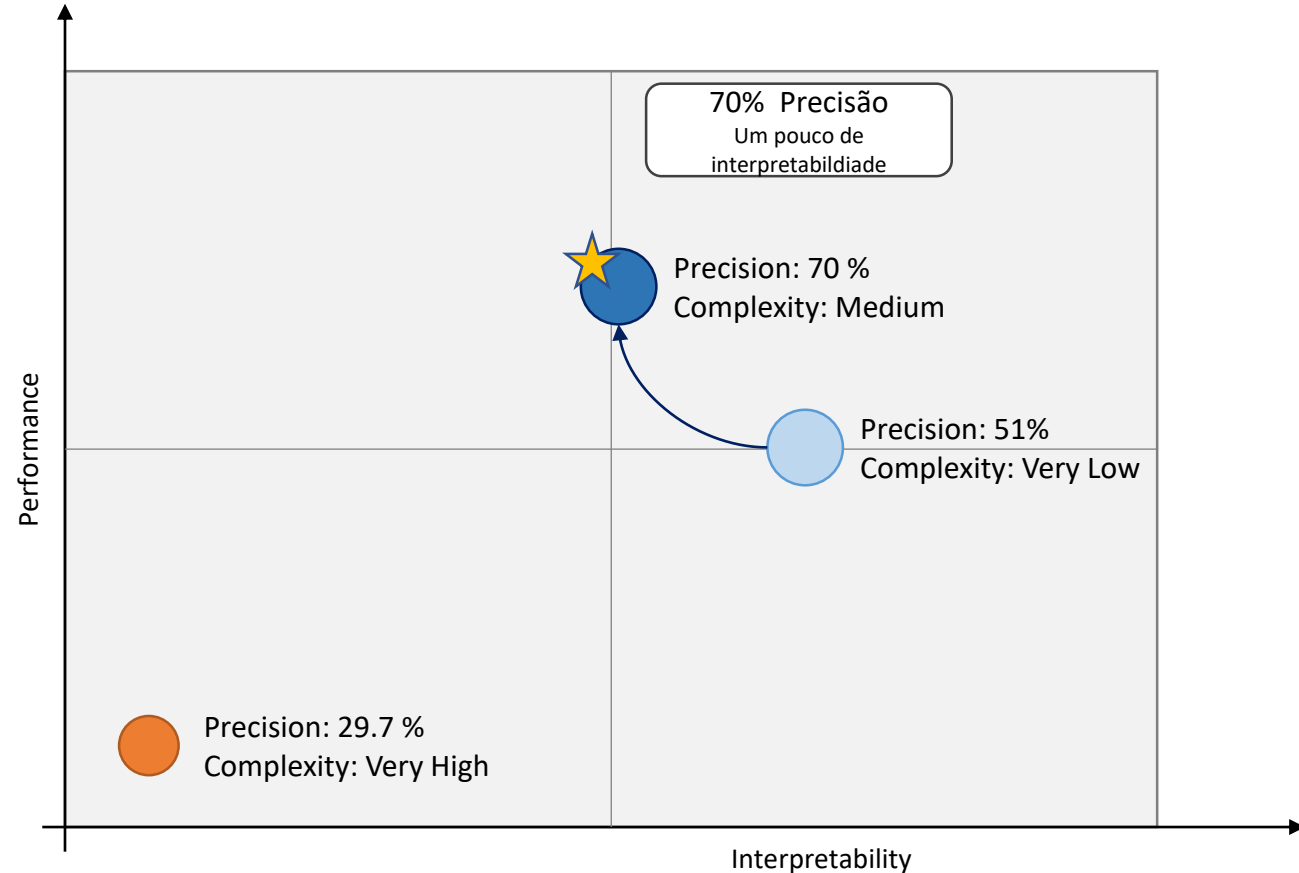
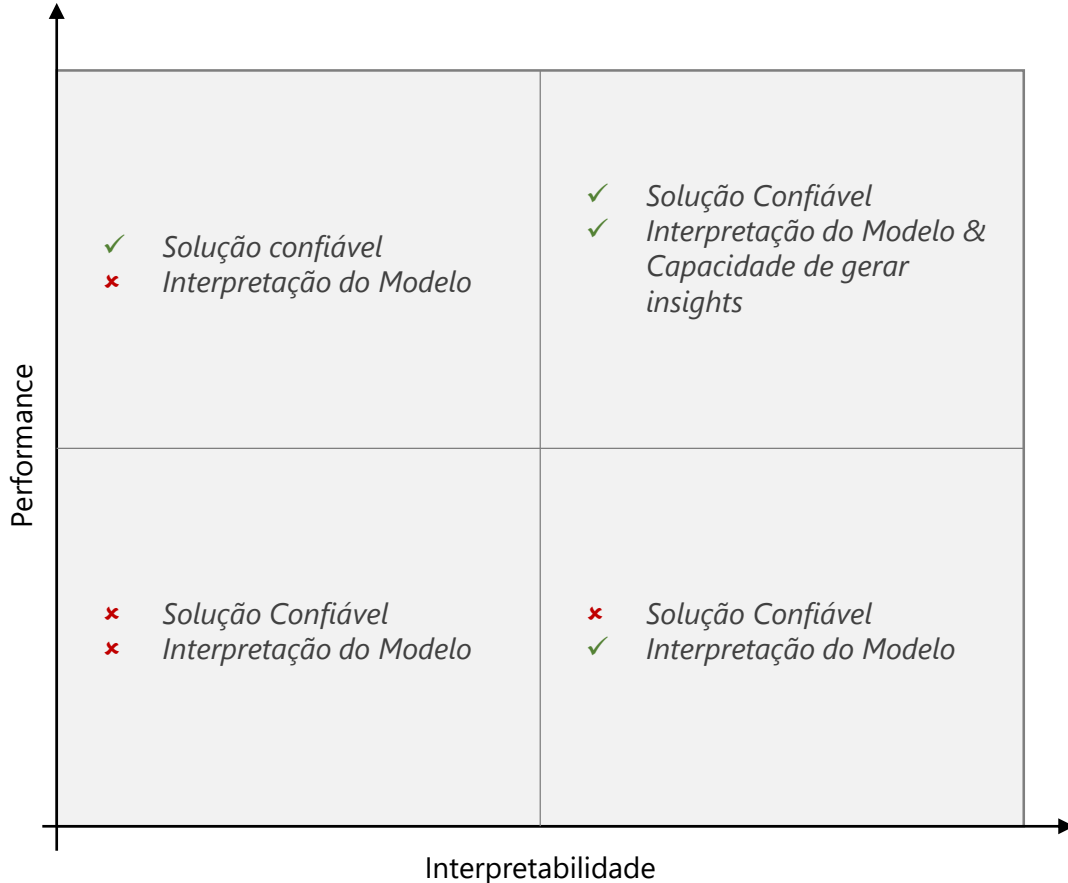
Modelo criado pela Elo Group



- **Técnica Utilizada:** Random Forest
- **Aumento de Precisão:** 71%
- **Direcionamento/Seleção** de uso de 10 variáveis
 - Definição de variáveis mais significativas

Precisão
0,71

Escolha dos Modelos: Interpretabilidade x Precisão



Performance:

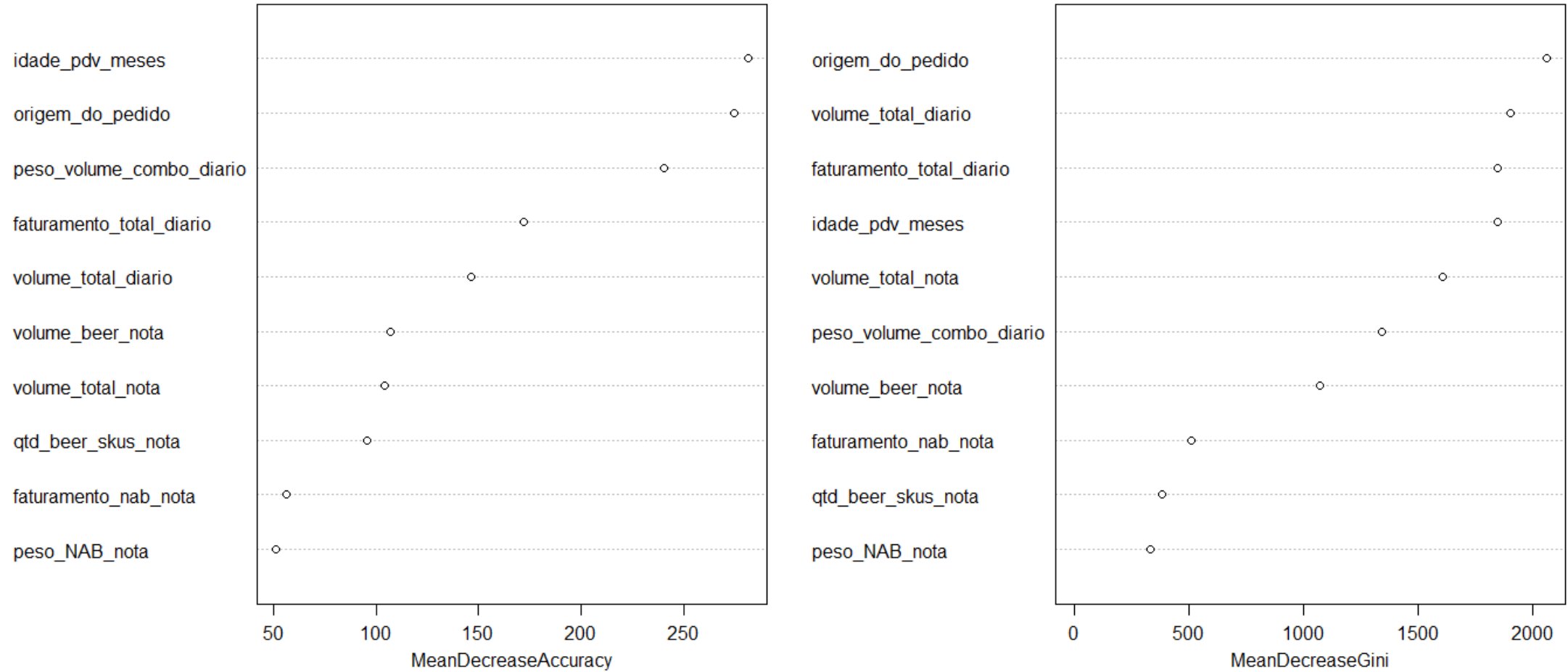
Foco do projeto em precisão para uma maior assertividade em casos de "Ponte"

Complexidade:

Como conseguimos interpretar o Algoritmo e seus resultados

VARIÁVEIS UTILIZADAS E SEU IMPACTO NO MODELO

modelo



DESENVOLVIMENTO



- Última Versão Disponível
- Modelo em Desenvolvimento

HOMOLOGAÇÃO

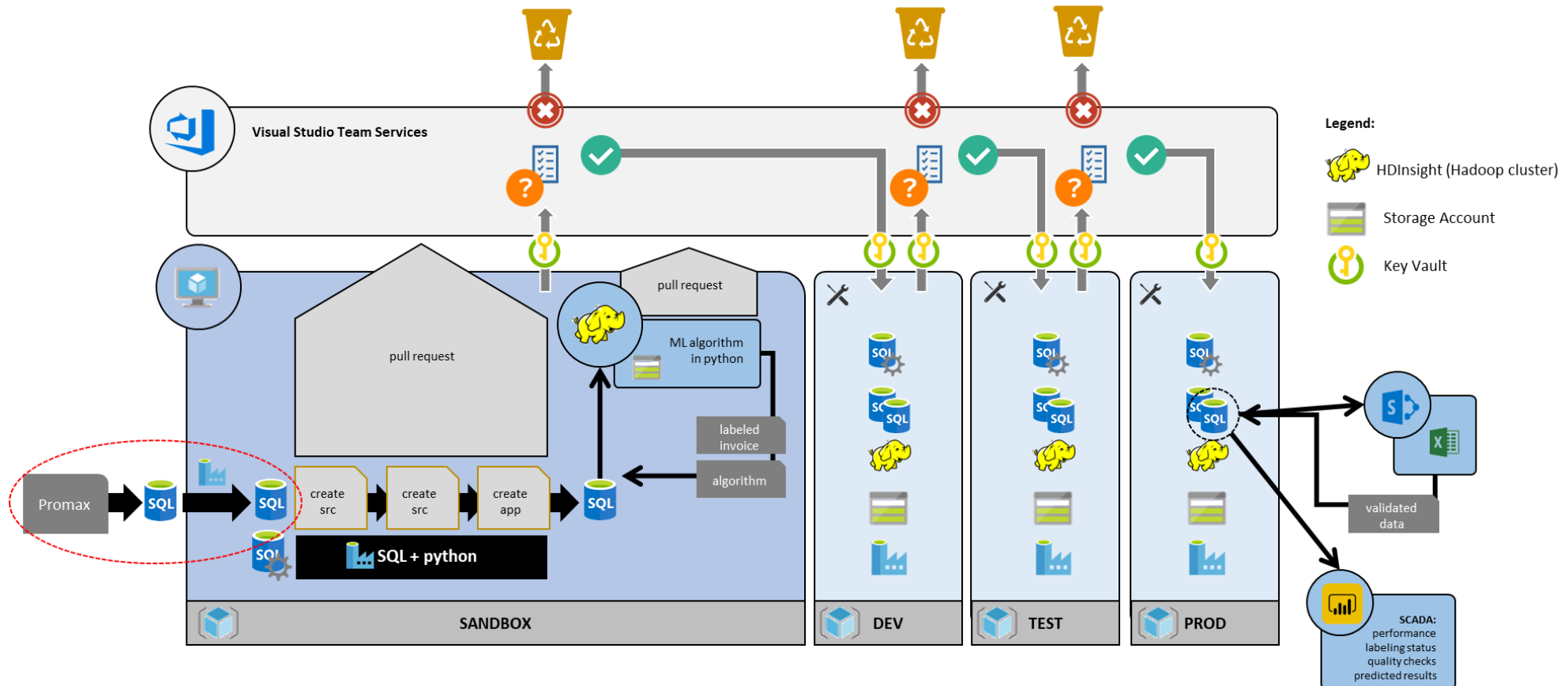


- Versão para Homologação

Produção



- Processo em Produção



Requirements: Python and R programming, basic knowledge of Hadoop commands & Hadoop architecture, basic Linux shell scripting



DÚVIDAS?

www.elogroup.com.br



OBRIGADO!

www.elogroup.com.br

FRANCISCO ALVES

21 99781-6456

Francisco.alves@elogroup.com.br