

# Detecção de Fraudes em Transações Financeiras



Desafio Técnico – Cientista de Dados  
Ana Carolina Costa Lourenço  
Junho/2025

# Objetivo

Desenvolver modelo para identificar fraudes em transações financeiras.

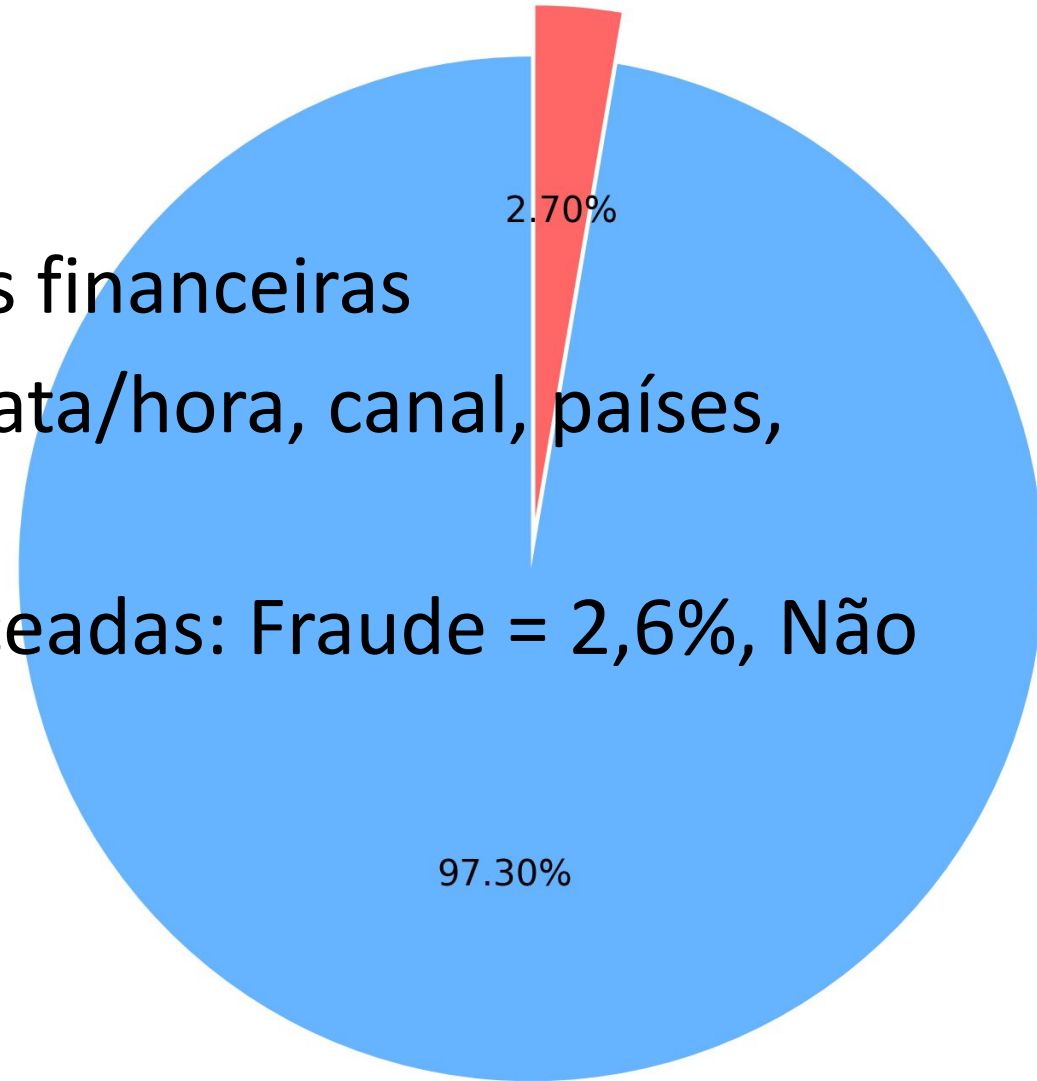
- Análise Exploratória
- Engenharia de Features
- Modelagem (ML)
- Avaliação de Desempenho

Entregáveis: Código, relatório técnico e recomendações.

# Base de Dados

- 10.200 transações financeiras
- Variáveis: valor, data/hora, canal, países, categoria, fraude
- Classes desbalanceadas: Fraude = 2,6%, Não Fraude = 97,4%

Proporção de Fraudes vs. Não Fraudes



Não Fraude

# Engenharia de Atributos

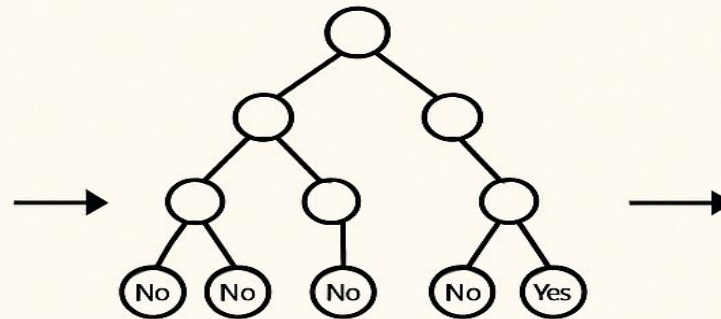
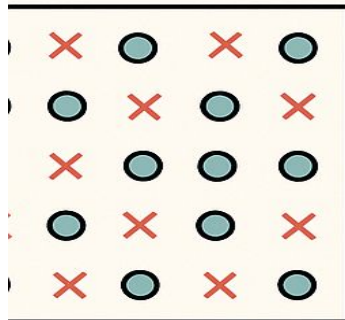
Criação de Novas Features:

- Nº Transações 24h
- Tempo desde última transação
- Valor médio histórico do cliente
- País de Origem  $\neq$  País de Destino

# Modelo Avaliado

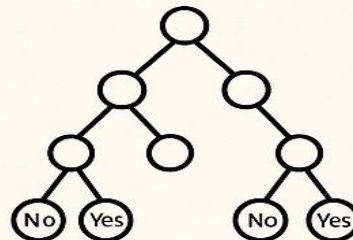
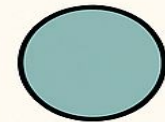
## Random Forest

Input Dataset



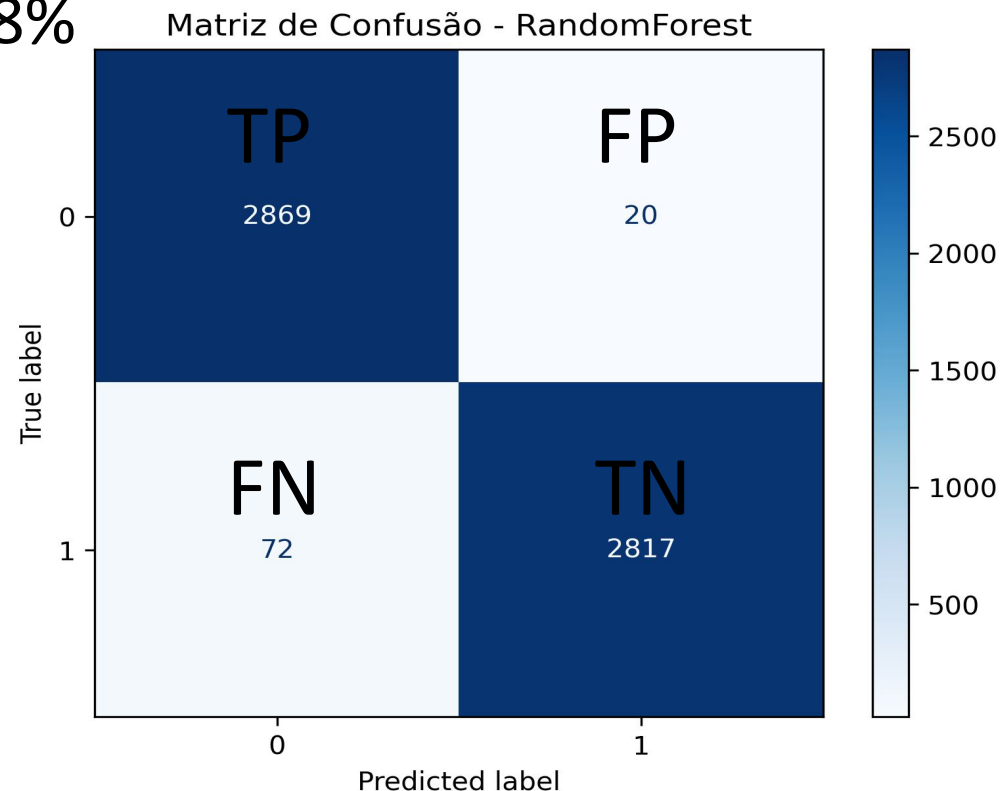
Tree 1 Tree 2 Tree 3

Prediction



# Performance do Random Forest

- AUC: 0.99 (ROC)
- Accuracy: 98,5%
- Recall (Fraude): 97,2%
- Precision (Fraude): 99,8%



# Análises Complementares

Fraudes mais frequentes quando:

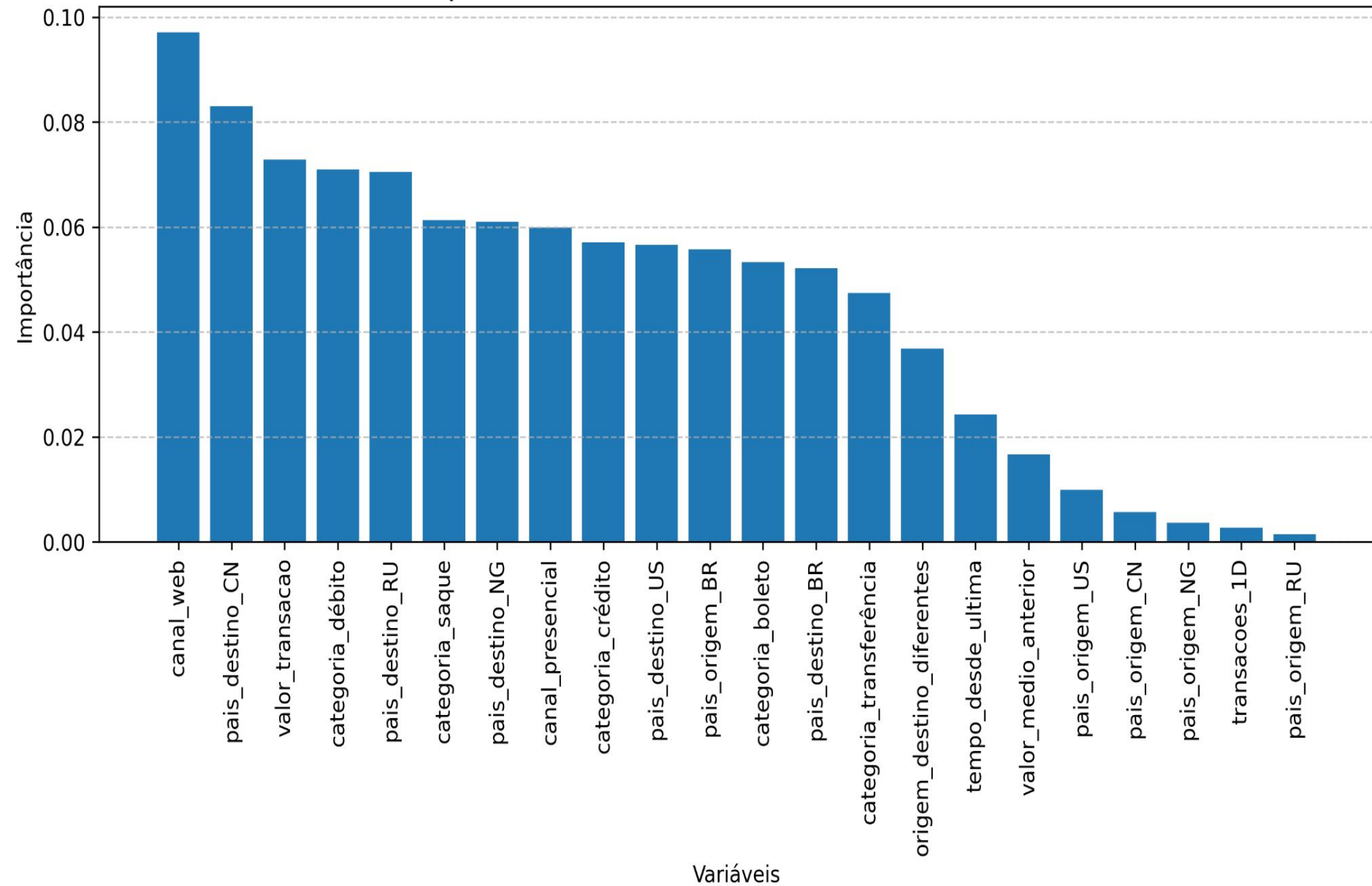
- Alta quantidade de transações 24h
- Pouco tempo desde última transação
- Transações internacionais

# Principais Achados

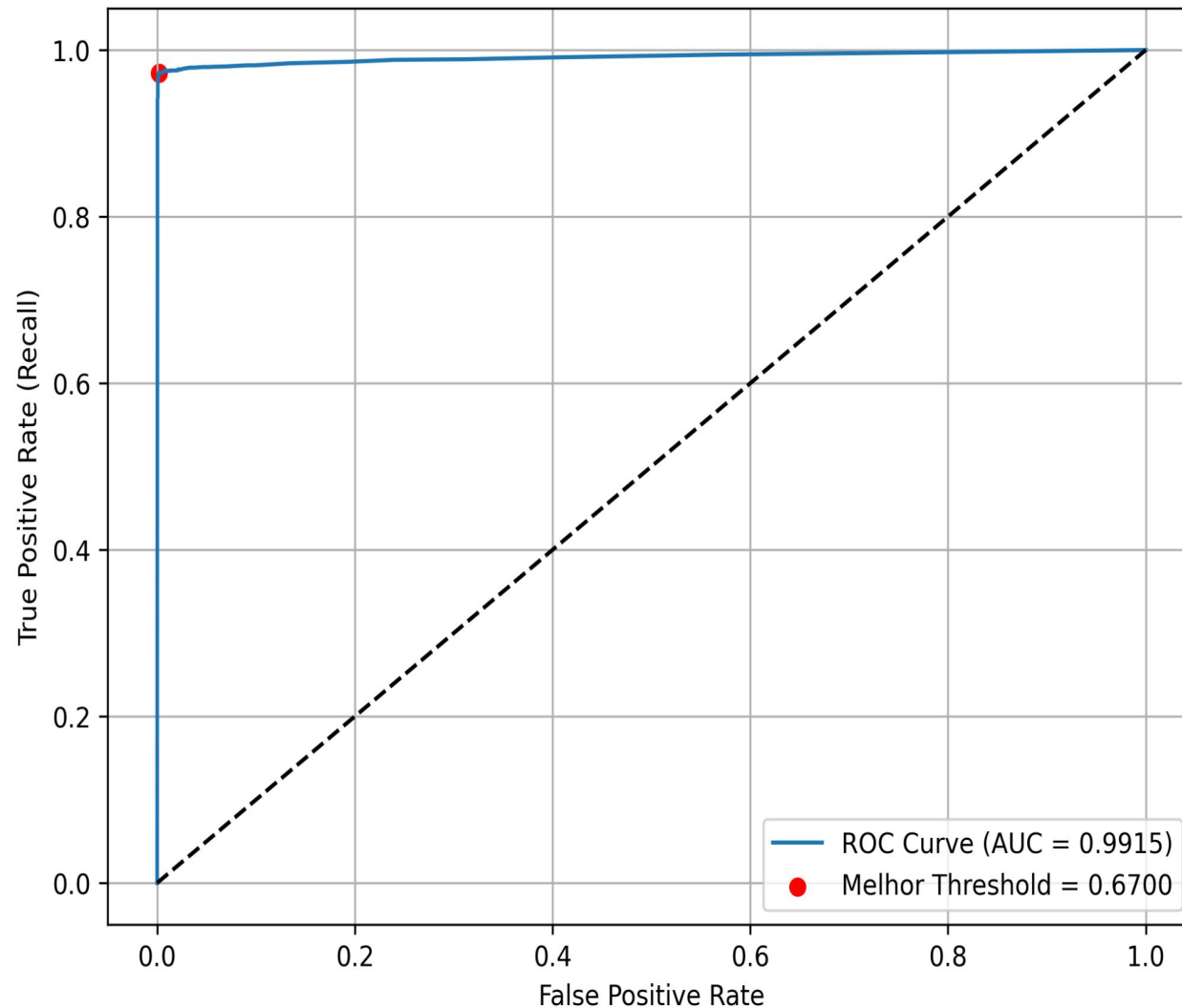
- ✓ Engenharia de atributos decisiva
- ✓ Modelo de random forest superou bastante o de KMeans
- ✓ Variáveis temporais e geográficas relevantes



# Importância das Variáveis - RandomForest



Curva ROC - RandomForest



Essa curva ROC em forma de L com  $AUC = 0.9$  pode está indicando que o modelo sofre com overfitting

# Recomendações

- Deploy do modelo Random Forest
- Monitoramento contínuo
- Análise com novos bancos de dados
- Integração com regras de negócio