



# Faixa de risco

Dados que salvam vidas

# Case

O projeto Faixa de Risco nasceu com o propósito de **transformar dados em ações que salvam vidas nas rodovias brasileiras**. Utilizando dados abertos da Polícia Rodoviária Federal (PRF) e tecnologias modernas de engenharia de dados e machine learning, construímos um pipeline completo e governado na AWS para prever a quantidade de acidentes de trânsito em BRs específicas.

Com uma acurácia de 90%, o modelo preditivo e os dashboards analíticos oferecem suporte estratégico à tomada de decisão da PRF, promovendo segurança viária com base em evidências. Mais do que um exercício técnico, este projeto reforça o poder dos dados como instrumento de impacto social.



Apenas em 2024...

**6K**

Mortos

**84K**

Feridos

**9.5M**

Infrações

# Integrantes



Carolina Althman

Cientista de dados



Guilherme Predolin

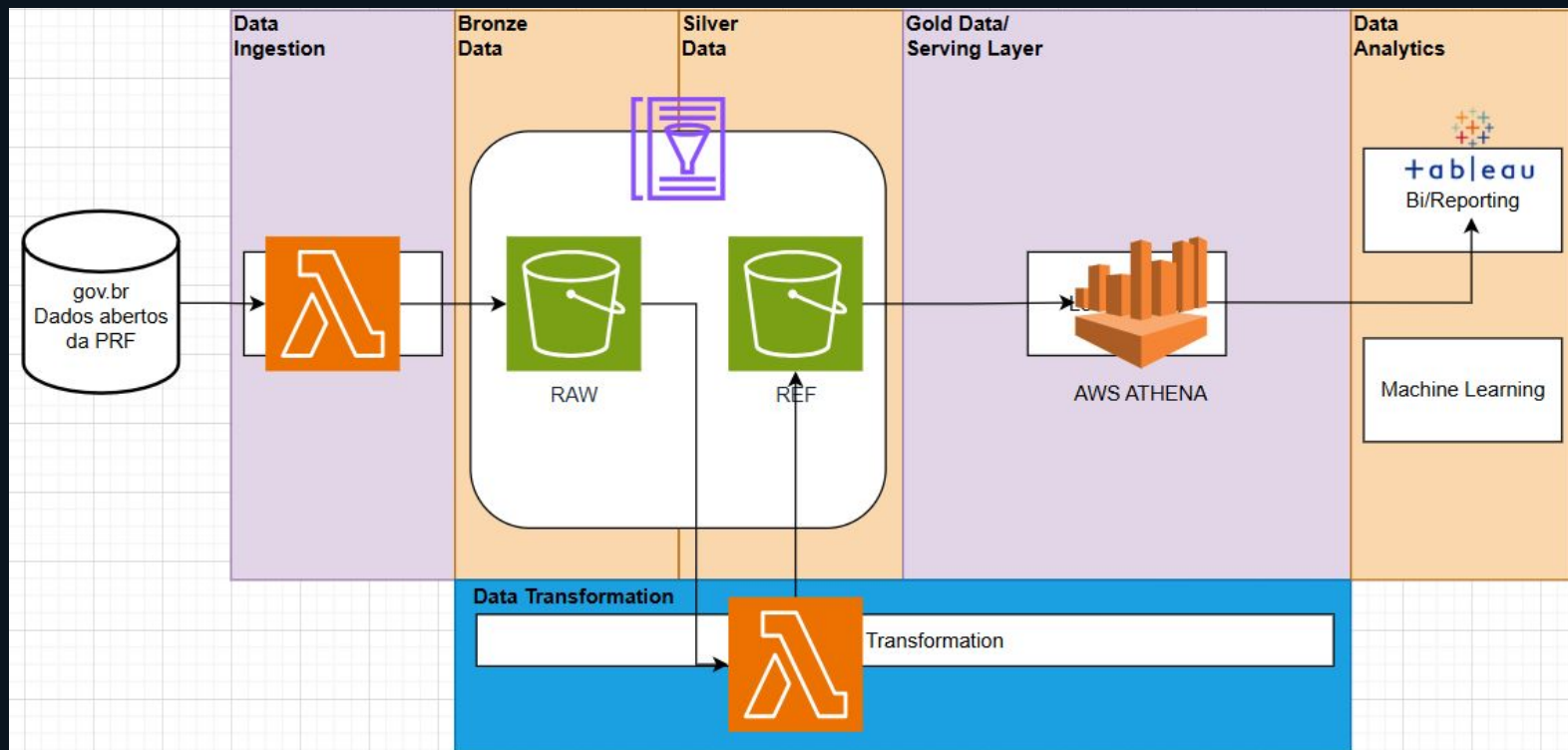
Engenheiro de dados



Italo Rufca

Analista de dados

# Arquitetura

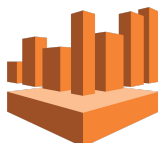


# Ferramentas

## Engenharia



Glue



Athena



S3



Lambda

## Visualização



Tableau

## Modelo



Pycaret



Colab

# Modelo de previsão de acidentes

Utilizamos o **PyCaret** para treinar diversos algoritmos e selecionar automaticamente o mais eficaz na previsão da quantidade de acidentes em rodovias federais.

**Dados utilizados para treino:** anos 2020 a 2022

**Dados utilizados para teste:** anos 2023 a 2025

O pipeline realiza:

- Ingestão dos dados já modelados
- Seleção automática do melhor modelo
- Predição de acidentes com base em variáveis como BR, KM, clima, feriados e histórico de infrações.

	Model	MAE	MSE	RMSE	R2	RMSLE	MAPE	TT (Sec)
et	Extra Trees Regressor	0.0576	0.0413	0.2028	0.8311	0.0682	0.0388	5.1899
lightgbm	Light Gradient Boosting Machine	0.0639	0.0436	0.2082	0.8219	0.0666	0.0404	7.5280
xgboost	Extreme Gradient Boosting	0.0685	0.0439	0.2091	0.8204	0.0686	0.0447	0.5230
gbr	Gradient Boosting Regressor	0.0878	0.0533	0.2305	0.7816	0.0755	0.0586	4.2360
rf	Random Forest Regressor	0.0794	0.0568	0.2379	0.7675	0.0791	0.0527	11.4860
dt	Decision Tree Regressor	0.0752	0.1078	0.3275	0.5588	0.1097	0.0482	0.3330
lar	Least Angle Regression	0.2270	0.2051	0.4527	0.1591	0.1539	0.1637	0.1760
br	Bayesian Ridge	0.2270	0.2051	0.4527	0.1591	0.1539	0.1637	0.1860
ridge	Ridge Regression	0.2270	0.2051	0.4527	0.1591	0.1539	0.1637	0.1820
lr	Linear Regression	0.2270	0.2051	0.4527	0.1591	0.1539	0.1637	0.5070
knn	K Neighbors Regressor	0.2051	0.2280	0.4773	0.0645	0.1646	0.1405	1.0560
en	Elastic Net	0.2487	0.2414	0.4910	0.0114	0.1637	0.1729	0.1720
lasso	Lasso Regression	0.2496	0.2429	0.4925	0.0056	0.1643	0.1735	0.2140
llar	Lasso Least Angle Regression	0.2496	0.2429	0.4925	0.0056	0.1643	0.1735	0.1800
omp	Orthogonal Matching Pursuit	0.2495	0.2431	0.4927	0.0046	0.1644	0.1735	0.2330
dummy	Dummy Regressor	0.2508	0.2443	0.4939	-0.0002	0.1648	0.1743	0.1590
par	Passive Aggressive Regressor	0.2389	0.2538	0.5036	-0.0420	0.1710	0.1619	0.2200
huber	Huber Regressor	0.1499	0.2796	0.5279	-0.1451	0.1747	0.0657	0.8030
ada	AdaBoost Regressor	0.5679	0.3899	0.6206	-0.5972	0.2543	0.5397	2.3800

# Resultado modelo

O modelo selecionado pelo AutoML foi o **Extra Trees Regressor**, com excelente desempenho na previsão da quantidade de acidentes por BR, KM e data.

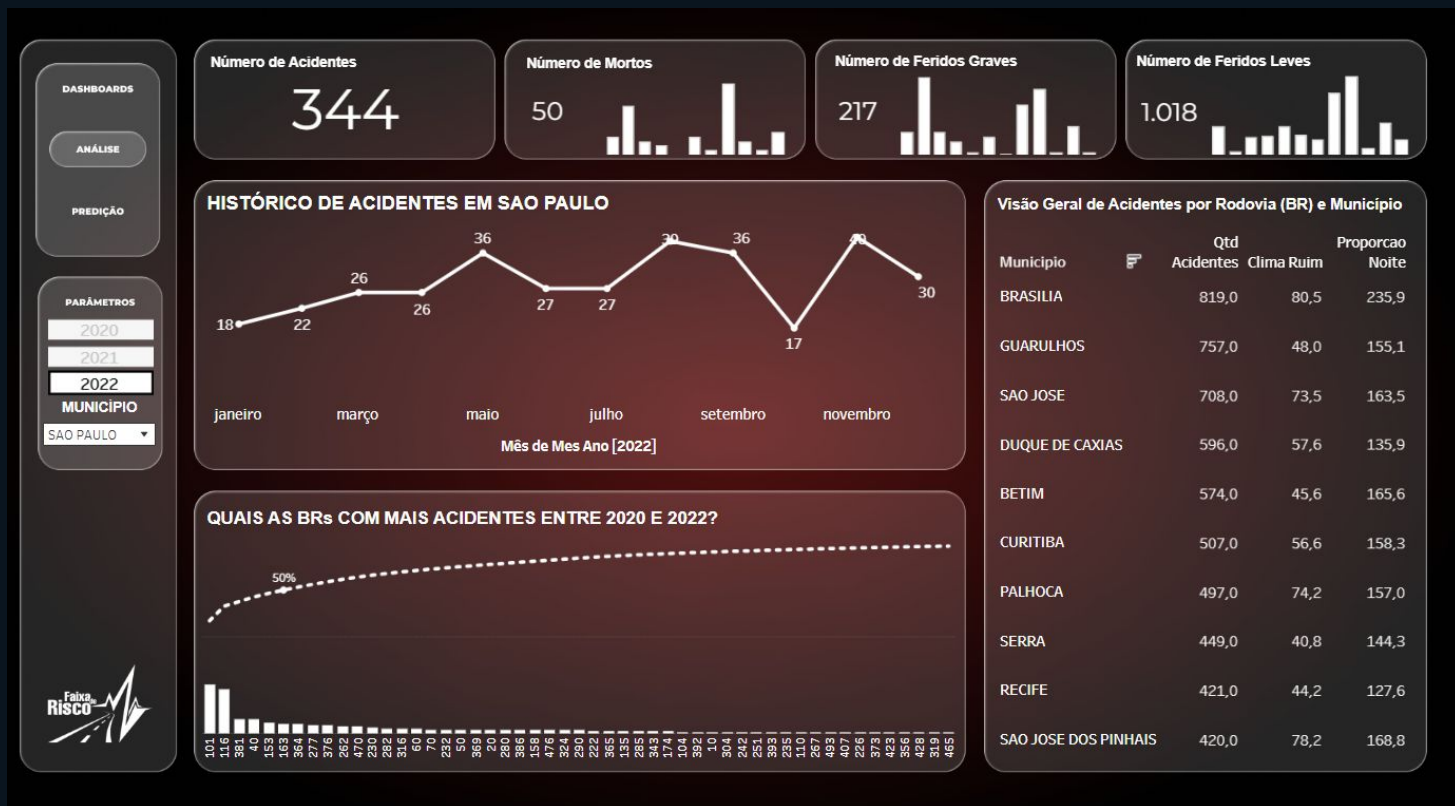
O modelo possui 90% de acerto nas previsões feitas.

Esses resultados comprovam a eficácia e robustez do modelo, tornando-o uma ferramenta confiável para apoiar decisões estratégicas de segurança viária e prevenção de acidentes.

Métrica	Descritivo da métrica	Valor
MAE	Média dos erros absolutos entre valores previstos e reais. Mede o erro médio direto.	0,0576
MSE	Média dos quadrados dos erros. Penaliza mais fortemente grandes erros.	0,0413
RMSE	Raiz quadrada do MSE. Representa o erro médio na mesma unidade da variável-alvo.	0,2028
R <sup>2</sup>	Proporção da variância explicada pelo modelo. Varia de 0 a 1 (quanto mais próximo de 1, melhor).	0,8311
RMSLE	Mede o erro em escala logarítmica, útil quando há grandes diferenças entre valores previstos e reais.	0,0682
MAPE	Média dos erros absolutos em relação ao valor real, em porcentagem.	0,0388
TT	Tempo necessário para treinar o modelo, em segundos.	6,1930



# Dashboard | Análise exploratória



# Dashboard | Previsão do modelo



### Parâmetros de Previsão

Data

2025/06/21

Município

ABADIA DE GOIAS

BR

60.0

Predizer 🚀

Desenvolvido por: Ítalo Rufca •  
Guilherme Predolin Dino • Carolina  
Althman

Fork

## Faixa de Risco

Plataforma para previsão de risco de acidentes e análise de dados.

Previsão

Análise Exploratória (EDA)

Modelagem (ML)

## Previsão de Risco de Acidentes

Aguardando parâmetros na barra lateral para gerar a previsão.



# Melhorias

Geração do modelo na ferramenta **sage maker**, pois possui integração direta com a arquitetura desenvolvida na AWS.

**Enriquecimento do dados em tempo real:** dados de clima, obras ou interdições, dados populacionais, etc.

Receber **feedback dos usuários** para identificar quais dados podem ser analisados para ajudar a evitar acidentes.

**As estatísticas de  
trânsito são feitas de  
pessoas que alguém  
amava.**

Obrigada!