

Previsão do Tempo de Sobrevivência de Pacientes Pós-Transplante

Análise Preditiva com Modelagem de Dados Clínicos

Desafio de Negócio

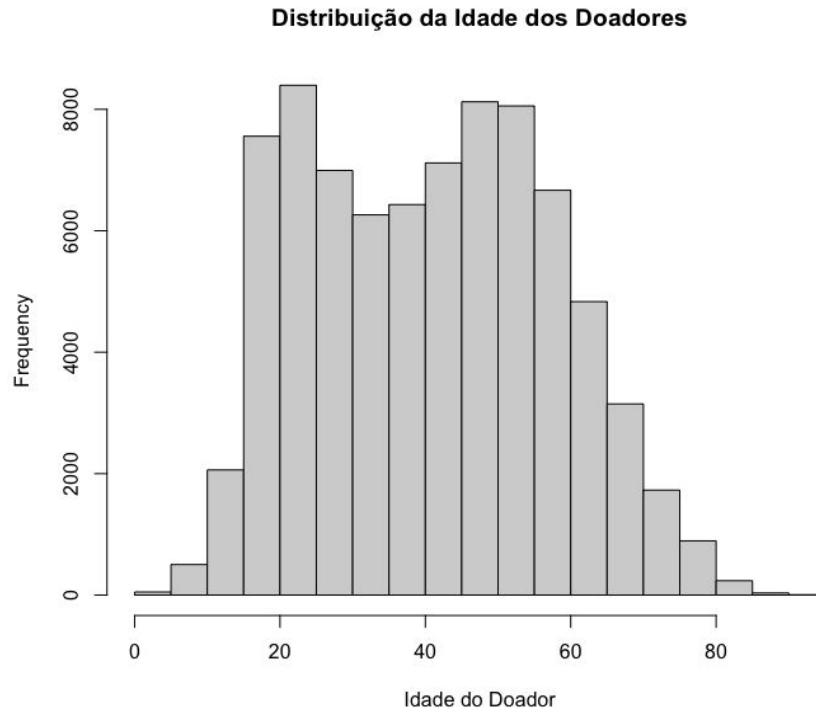
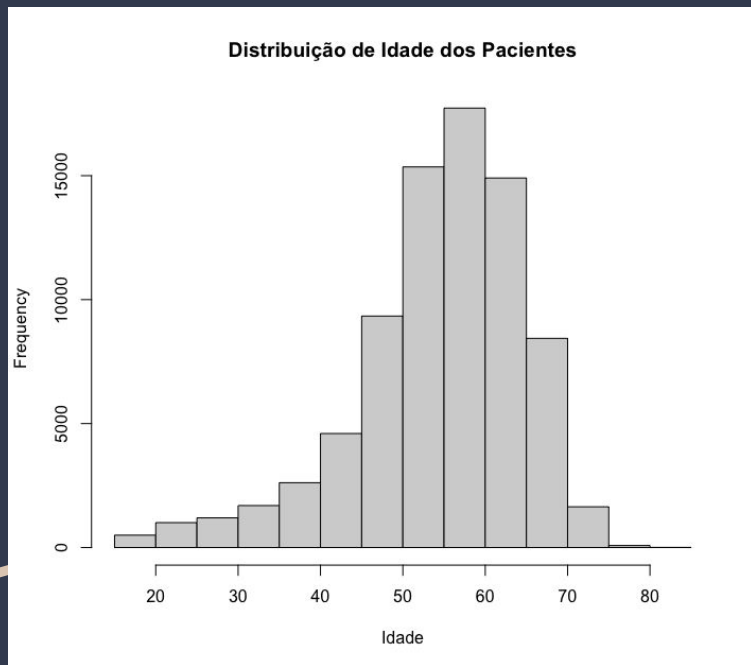
Precisamos prever o tempo de sobrevivência dos pacientes após o transplante.

Com objetivo de ajudar na personalização do tratamento e otimização do acompanhamento.



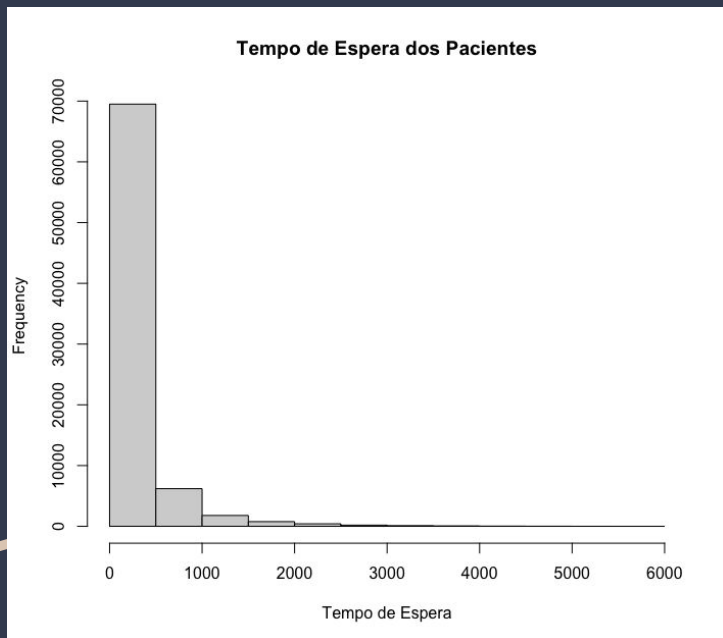
Análise Exploratória

Análise da distribuição de variáveis

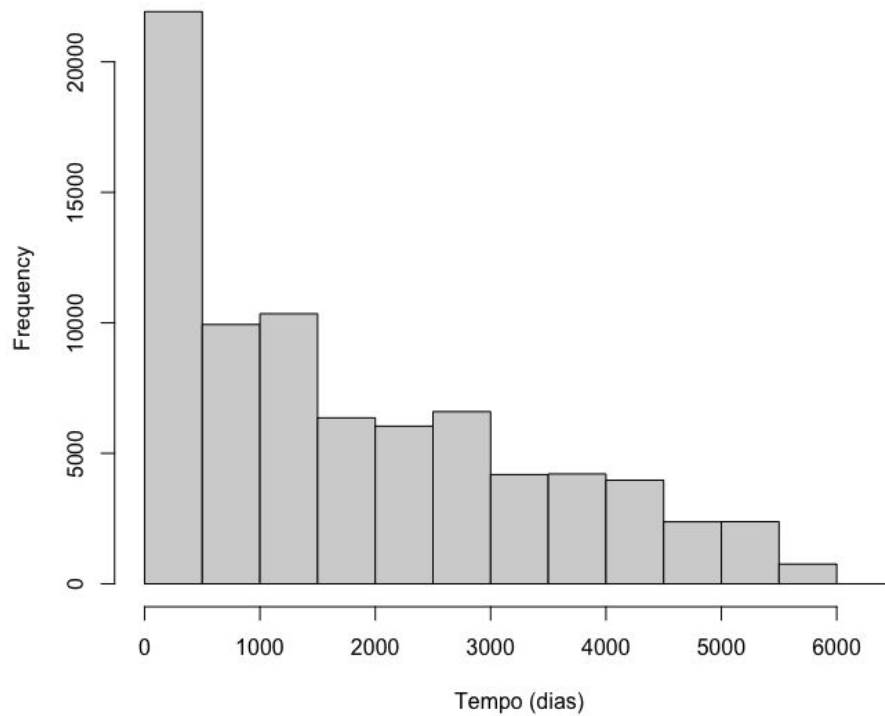


Análise Exploratória

Análise da distribuição de variáveis

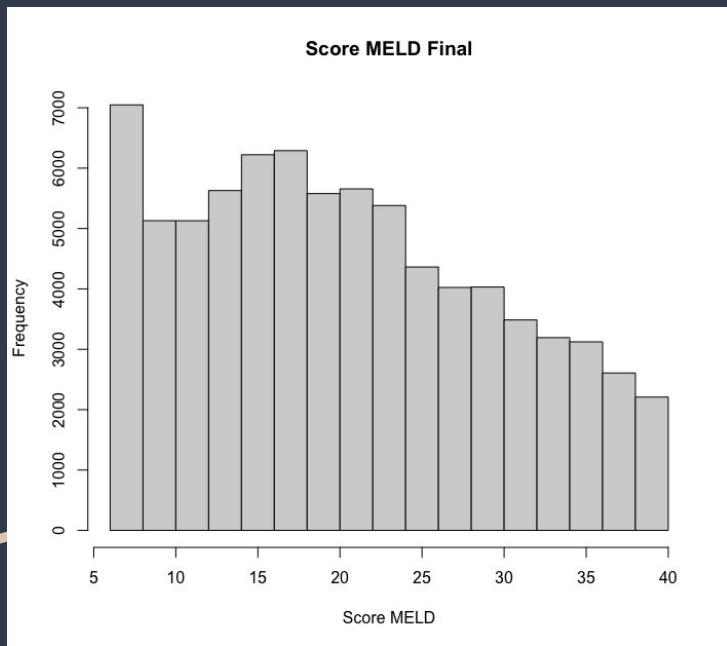


Tempo de Sobrevivência dos Pacientes

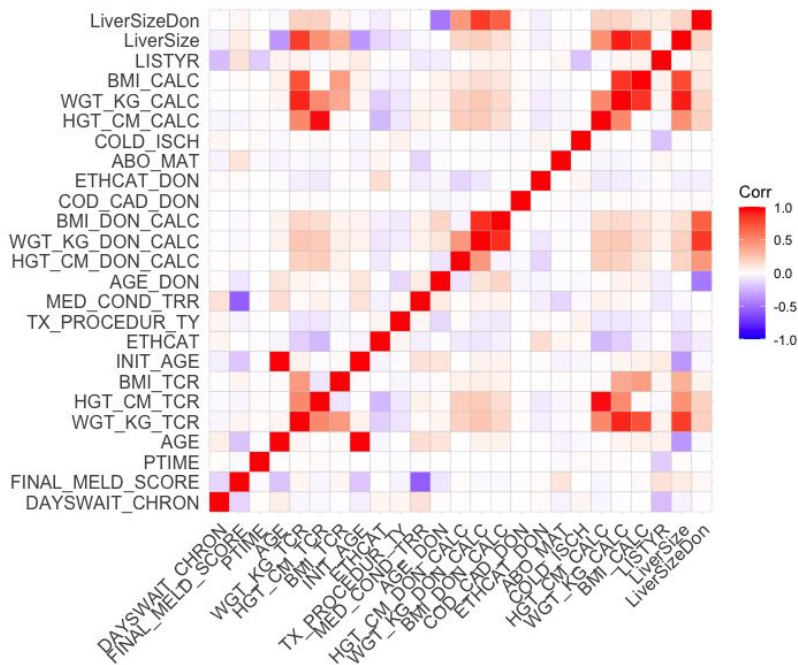


Análise Exploratória

Análise da distribuição de variáveis



Correlação entre as variáveis numéricas



Preparação para Modelagem

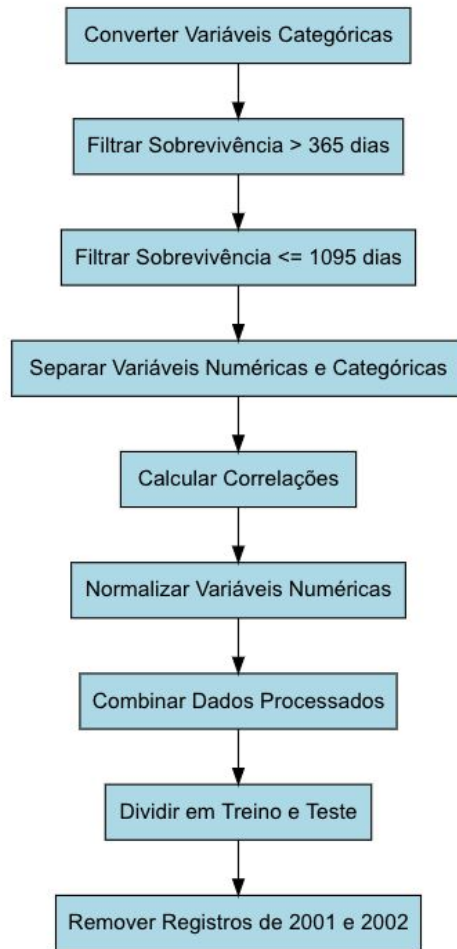
Filtragem de sobreviventes > 365 dias e <= 1095 dias

Conversão de variáveis categóricas em fatores

Padronização de variáveis numéricas

Separação entre variáveis numéricas e categóricas

Filtragem dos dados 2001 e 2002



Modelo de Regressão Linear

Predição do tempo de sobrevivência baseado em variáveis como

PTIME

FINAL_MELD_SCORE

REGION

AGE

GENDER

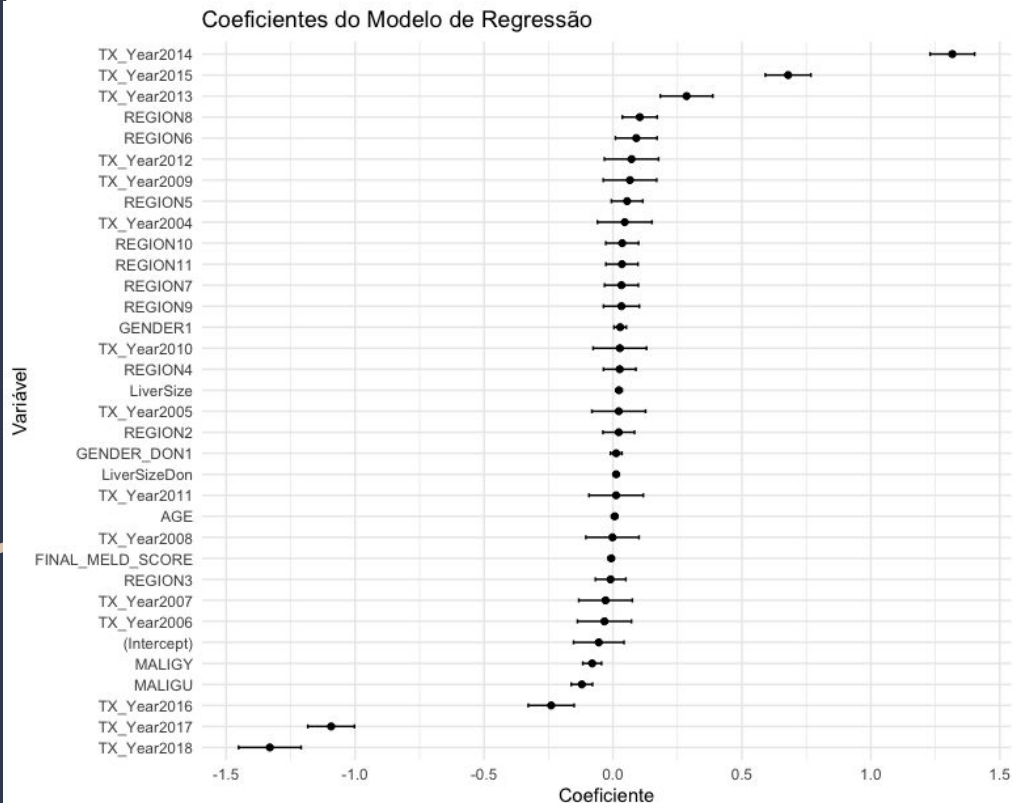
GENDER_DON

LiverSize

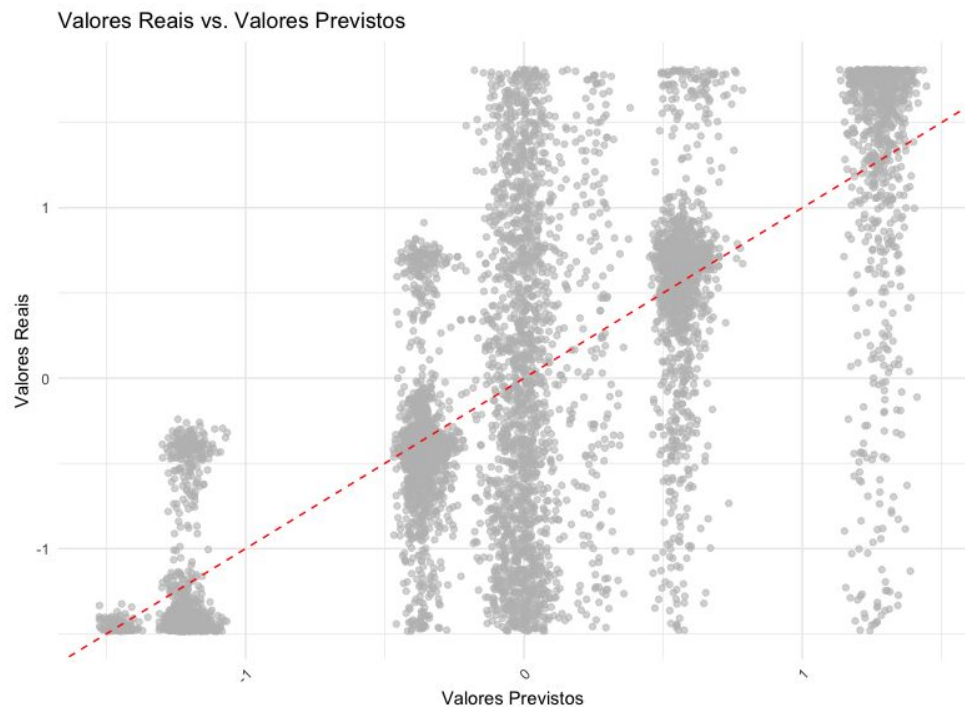
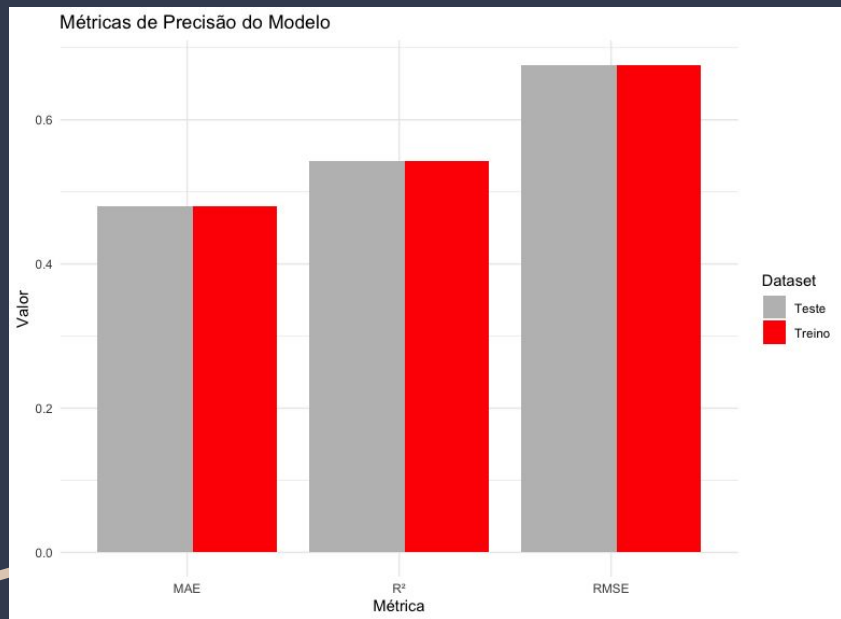
LiverSizeDon

MALIG

TX_Year



Resultados do modelo de regressão linear



Modelo de XGBoost

Etapas de Preparação dos Dados

Divisão dos dados (70%/30%)

Remoção dos anos de 2001 e 2002

Padronização das variáveis numéricas

Separação das variáveis numéricas e categóricas

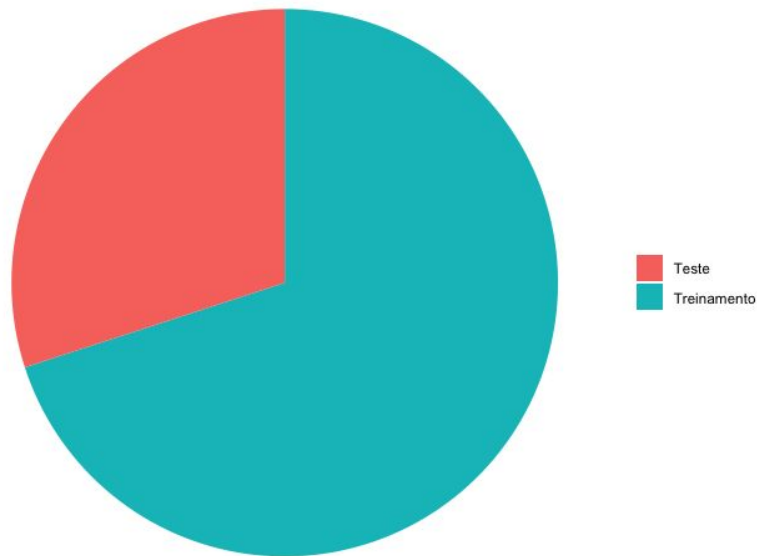
Parâmetros

Utilização do booster gbtree com regressão quadrática

Parâmetros ajustados

Early Stopping

Divisão dos Dados em Treino e Teste



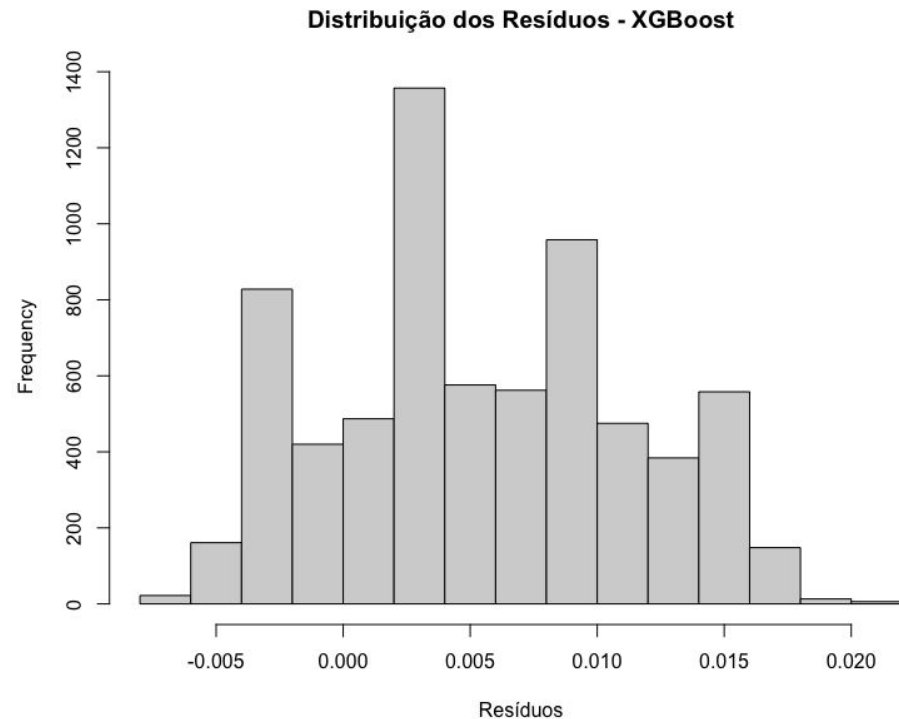
Modelo de XGBoost

Resultados - Treinamento / Teste

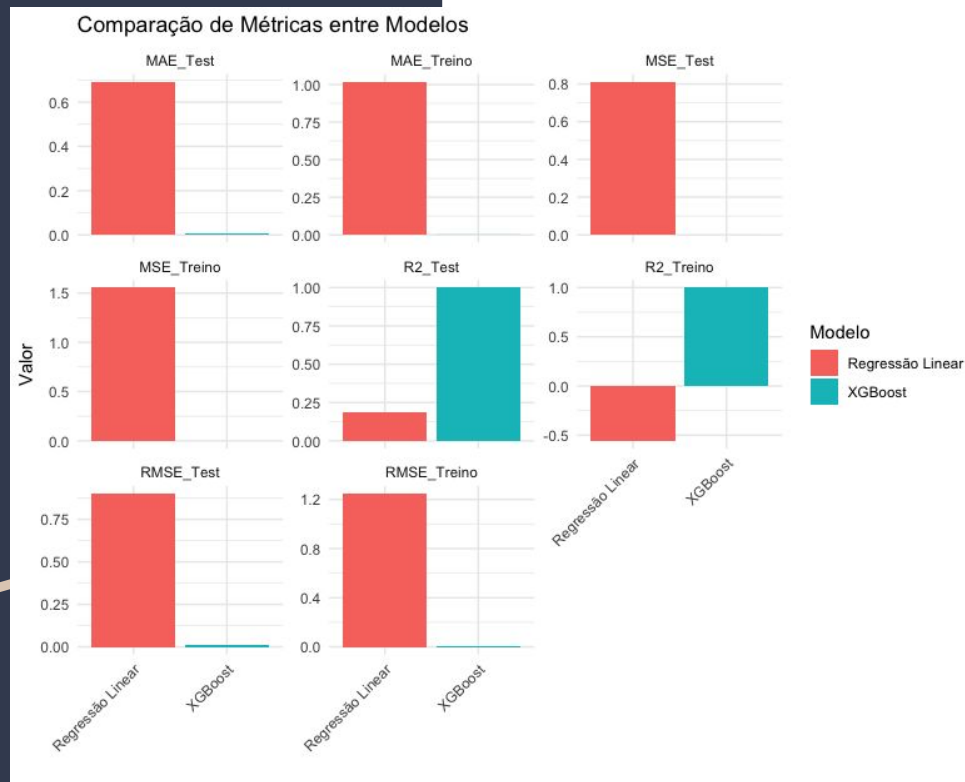
RMSE: 0.0068 / 0.0039

MAE: 0.0057 / 0.0031

R²: 0.99995 / 0.99998



Comparação de Métricas: Regressão Linear vs. XGBoost



Conclusões e Próximos Passos

O modelo XGBoost apresentou resultados superiores, com menor RMSE e MAE nos dados de teste

O R^2 próximo de 1 para o XGBoost indica uma previsão muito precisa

A regressão linear, embora eficaz, mostrou um desempenho inferior quando comparado ao XGBoost

Com o modelo escolhido, agora podemos colocar em produção em um ambiente real para otimização de processos

