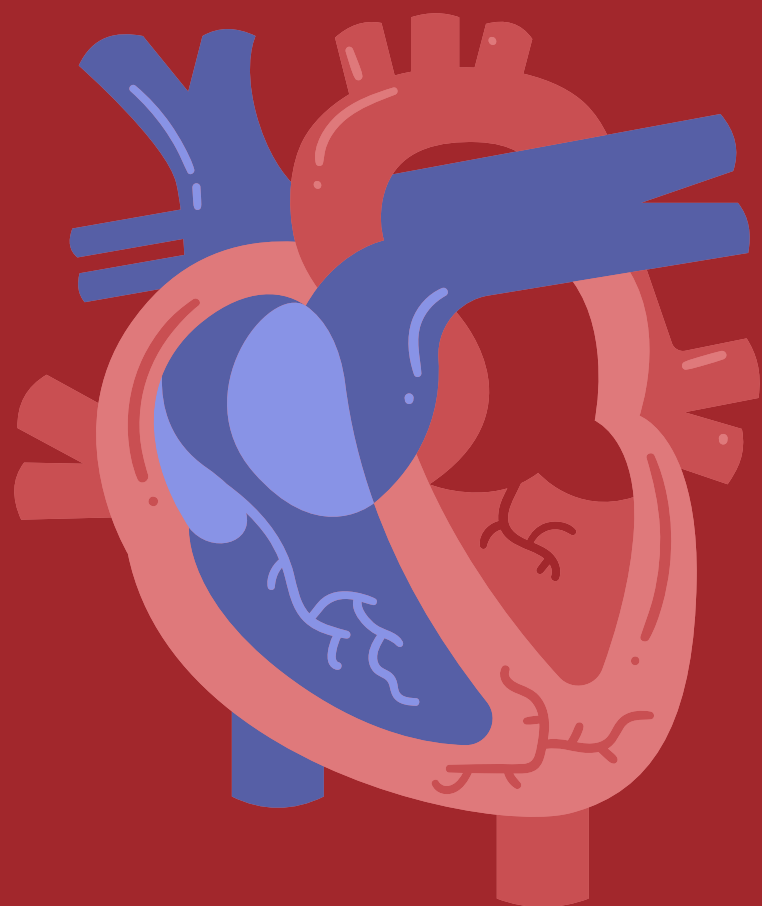


IA PARA PREDIÇÃO DE DOENÇA CARDIACA



**INTEGRANTES:
MARIANA FURTADO
FELIPE CARVALHO**

SUMÁRIO:

1. Motivação
2. Dataset
3. Metodologia
4. Análise dos Resultados



MOTIVAÇÃO

- As doenças cardiovasculares (DCV) são a principal causa de mortalidade global
- Modelos podem analisar um conjunto de dados, facilitando a detecção precoce e a gestão eficaz de indivíduos em risco
- Desenvolvimento de ferramentas preditivas não só aprimora a precisão dos diagnósticos, mas também alivia a carga sobre os sistemas de saúde,

DATASET

12 colunas - 70.000 linhas

id: Identificação única do indivíduo.

age: Idade do indivíduo em dias.

height: Altura do indivíduo em centímetros.

weight: Peso do indivíduo em quilogramas.

gender: Gênero do indivíduo, representado por valores categóricos.

ap_hi: Pressão arterial sistólica medida em milímetros de mercúrio (mmHg).

ap_lo: Pressão arterial diastólica medida em milímetros de mercúrio (mmHg).

cholesterol: Nível de colesterol, classificado em três categorias.

gluc: Nível de glicose.

smoke: Indica se o indivíduo é fumante.

alco: Indica se o indivíduo consome álcool.

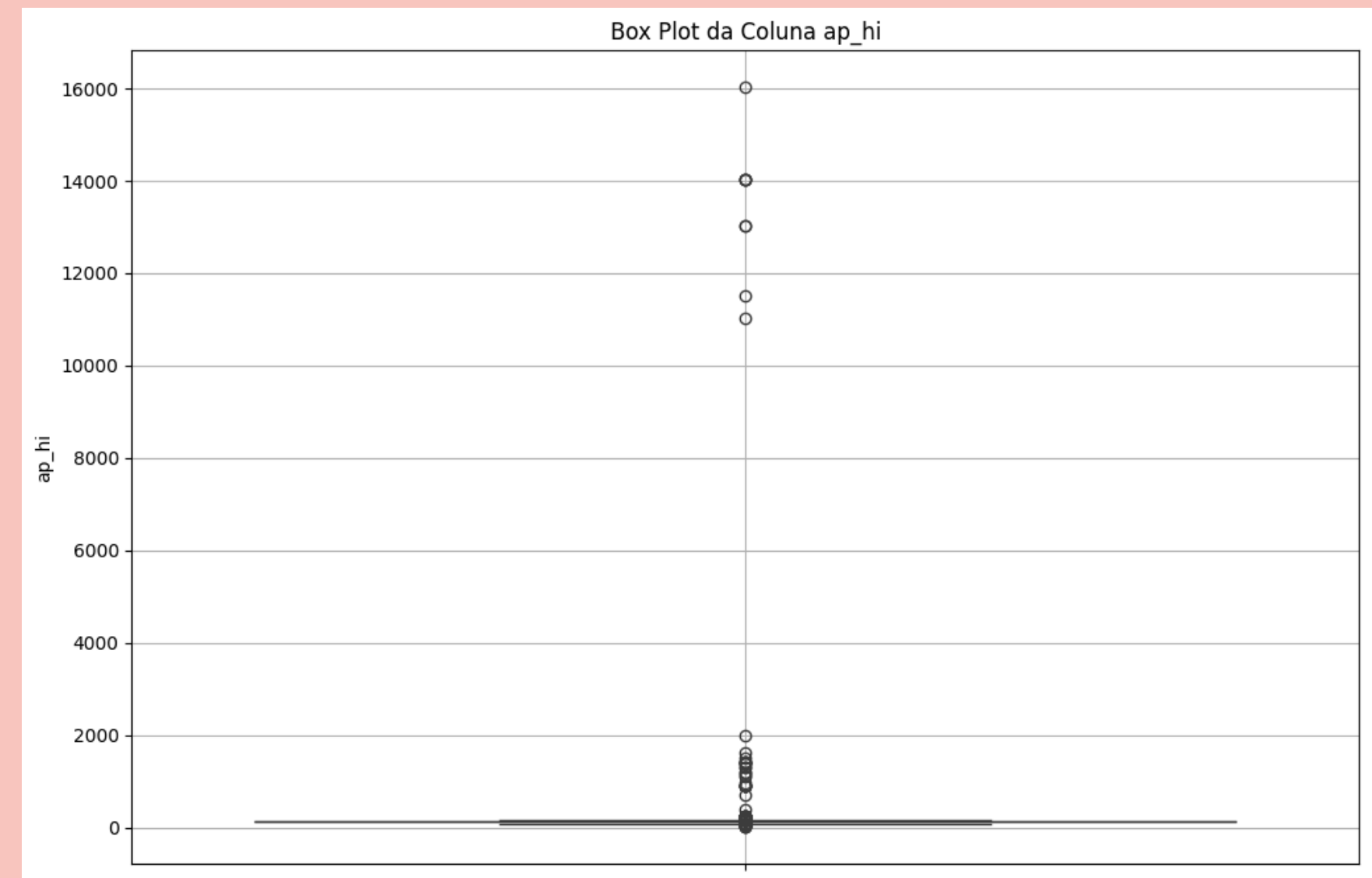
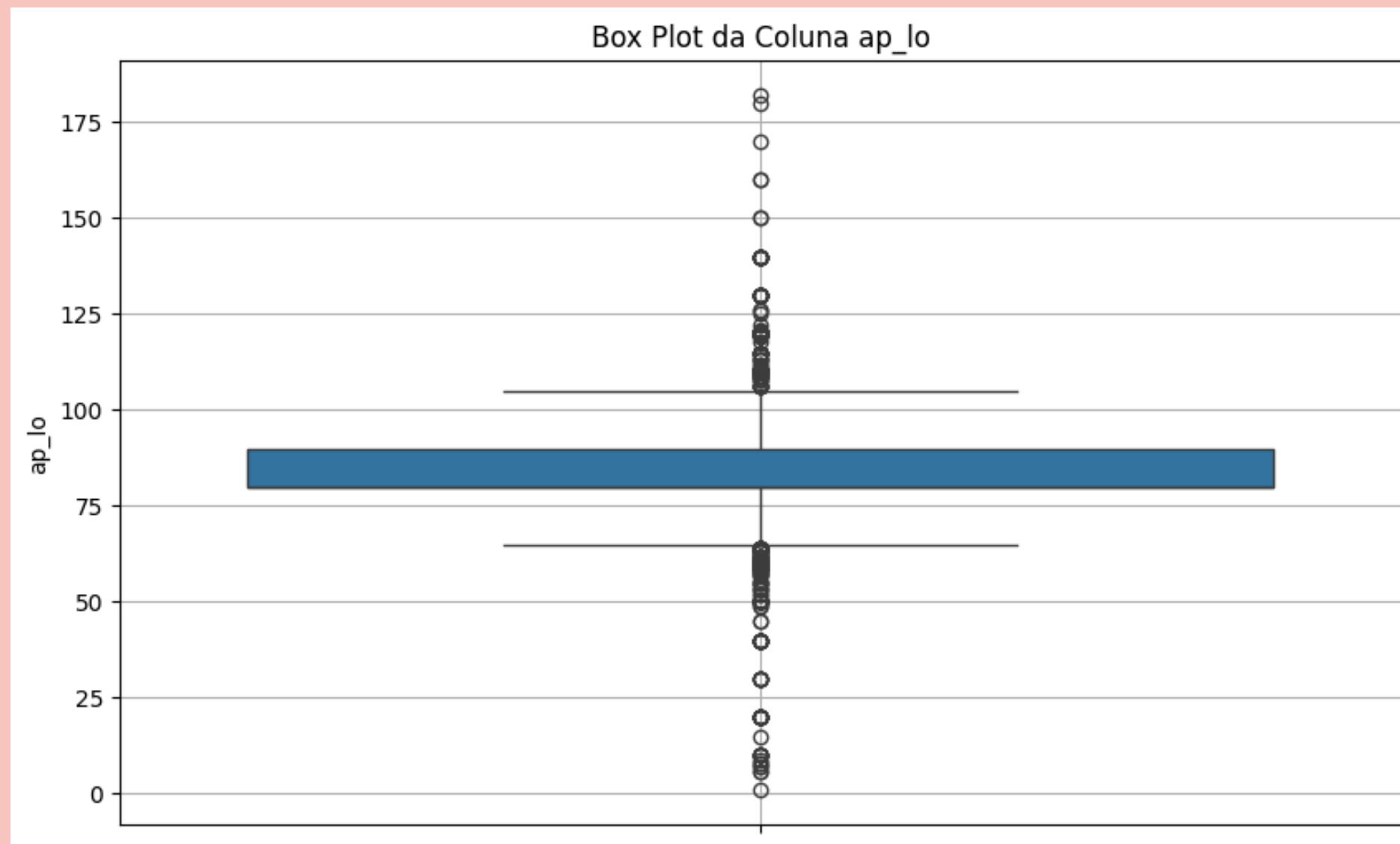
active: Indica se o indivíduo pratica atividade física regularmente.

cardio: Indica a presença de doença cardiovascular. 

METODOLOGIA

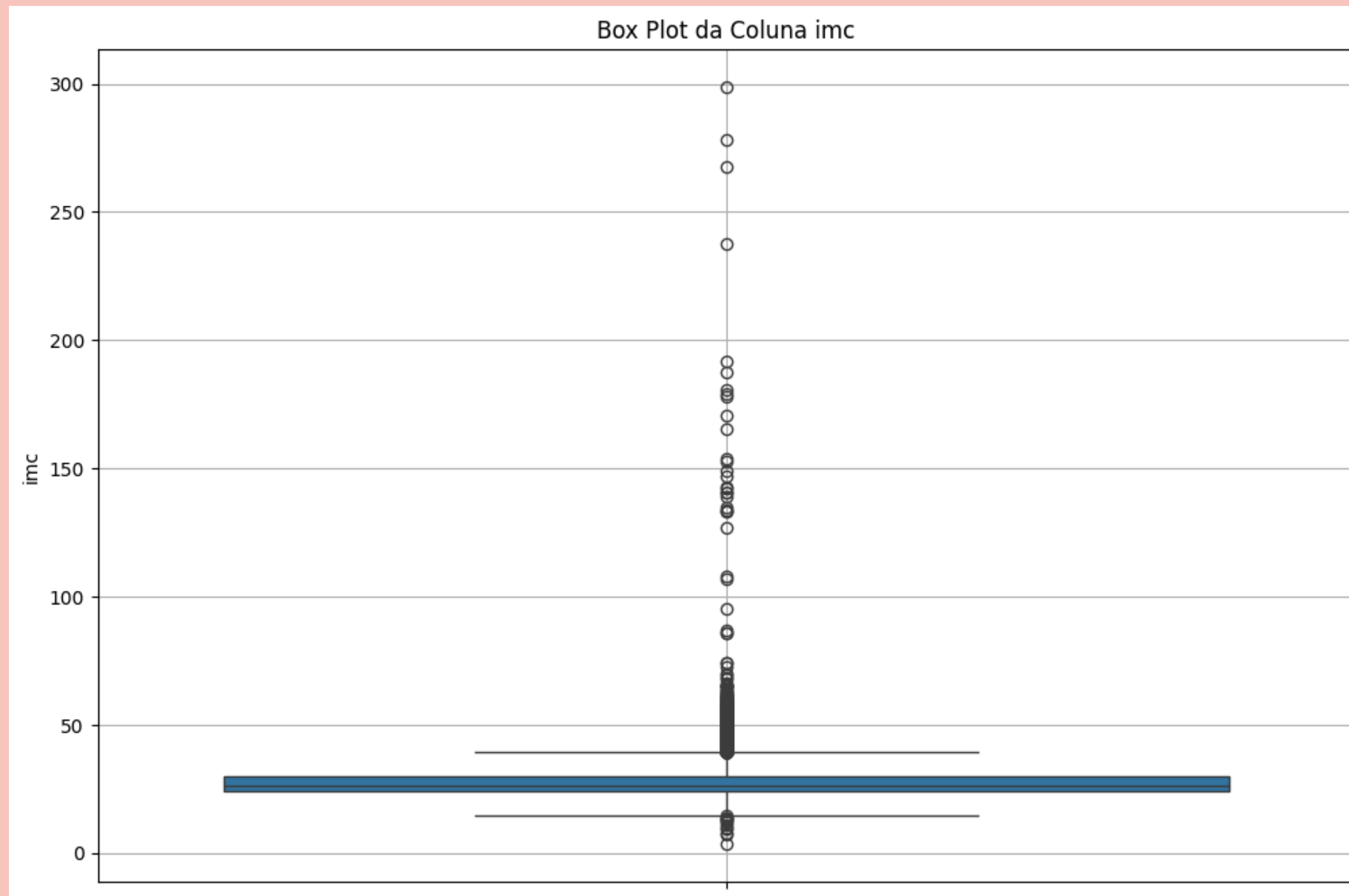
Pré-processamento dos Dados

Limpeza dos Dados: Foram removidos registros com valores ausentes, duplicados ou inconsistentes, como pressões arteriais inválidas.



Pré-processamento dos Dados

Transformação dos Dados: As variáveis categóricas foram codificadas usando One-Hot Encoding, e os dados numéricos foram padronizados usando a técnica StandardScaler. A coluna 'age' foi transformada em anos, além disso, foi feita uma coluna 'imc' com as variáveis 'weight' e 'height'



Modelos:

- **Árvore de decisão**
- **Random Forest**
- **Rede Neural**

Os modelos foram ajustados com ajuda da RandomizedSearchCV para identificar os melhores hiperparâmetros.

Foi utilizada a validação cruzada com 5 folds (KFold com n_splits=5).

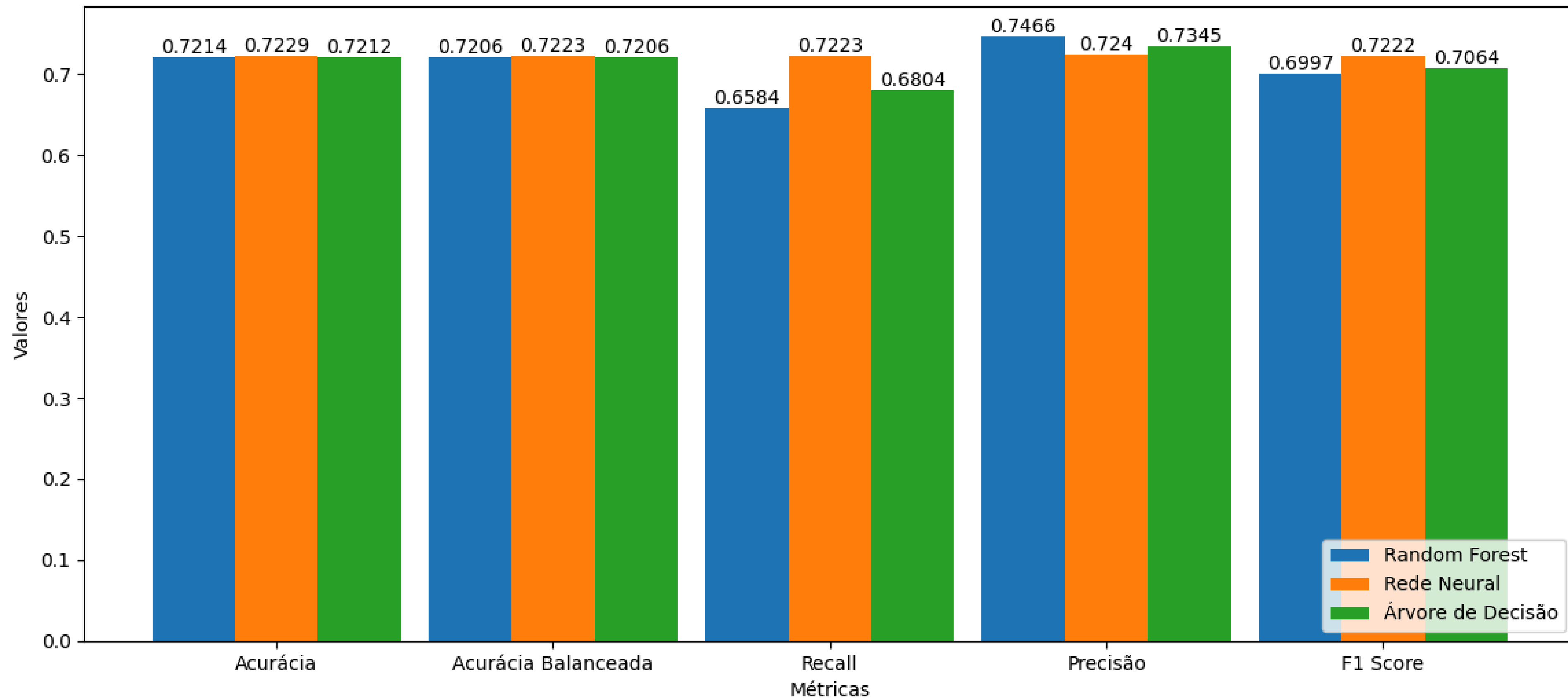
Os modelos foram ajustado com os melhores hiperparâmetros identificados na etapa anterior.

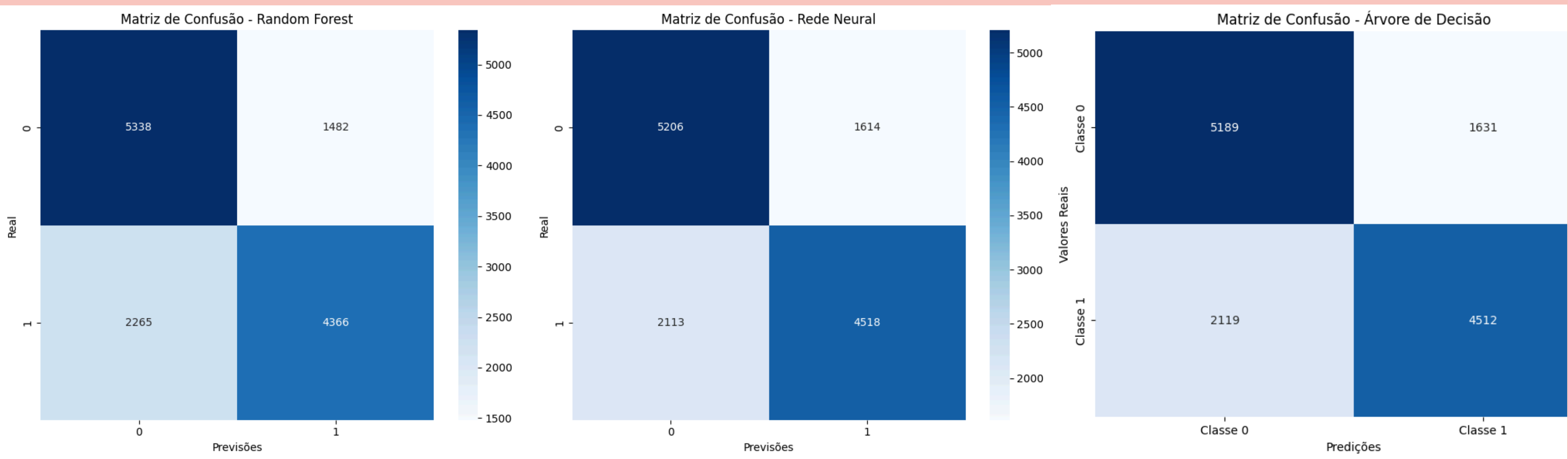
Após a validação cruzada, os modelos finais foram treinados com todos os dados de treino e avaliados no conjunto de teste. As métricas finais incluíram:

- Matriz de Confusão
- Acurácia
- Acurácia Balanceada
- Recall
- Precision
- F1-Score

ANÁLISE DOS RESULTADOS

Comparação de Métricas entre Random Forest, Rede Neural e Árvore de Decisão





Obrigado

