

# Predição de Mortalidade em Insuficiência Cardíaca

Ítalo Rosa Gonçalves

Instituto Politécnico do Rio de Janeiro (IPRJ)  
Universidade do Estado do Rio de Janeiro (UERJ)

Dezembro de 2025



IPRJ  
Universidade do Estado  
do Rio de Janeiro



# Contexto e Definição do Problema

- **Objeto de Estudo:** Pacientes com Insuficiência Cardíaca.
- **Dataset:** *Heart Failure Clinical Records* (299 registros).
- **Objetivo:** Predizer a variável alvo DEATH\_EVENT.
- **Desbalanceamento:**
  - Classe 0 (Sobreviventes):  $\approx 68\%$
  - Classe 1 (Óbitos):  $\approx 32\%$
- **Problema Clínico:** Modelos tradicionais tendem a priorizar a classe majoritária, gerando Falsos Negativos nos pacientes de risco.

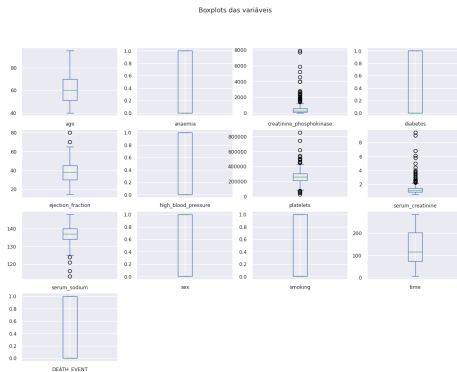


Figure: Distribuição das Variáveis Clínicas

# Bloxplot das Features

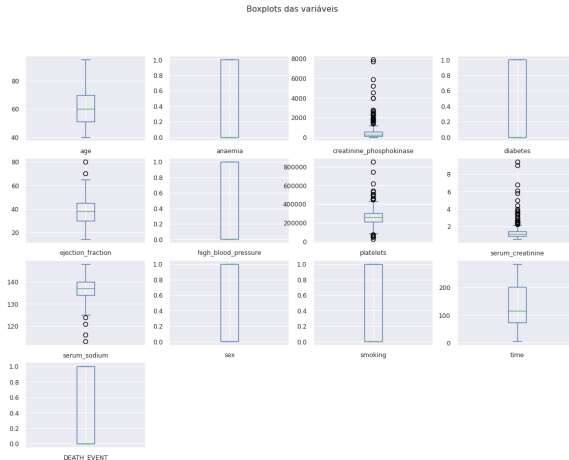


Figure: Divisão das variáveis

A abordagem foi estruturada em cinco etapas fundamentais:

- ❶ **Pré-processamento:** Aplicação de *StandardScaler* para normalização das escalas.
- ❷ **Otimização de Hiperparâmetros:** Utilização do algoritmo de **Evolução Diferencial** para busca global.
- ❸ **Modelagem:** Comparação entre *Random Forest* (Ensemble) e *Multilayer Perceptron* (Rede Neural).
- ❹ **Validação:** *K-Fold Cross-Validation* ( $k = 5$ ) com 20 execuções independentes.

A abordagem foi estruturada no mesmo molde do experimento 1:

- 1 **Balanceamento (ADASYN):** Geração adaptativa de amostras sintéticas apenas no conjunto de treino para reforçar a classe minoritária.
- 2 **Otimização de Hiperparâmetros:** Utilização do algoritmo de **Evolução Diferencial** para busca global.
- 3 **Modelagem:** Comparação entre *Random Forest* (Ensemble) e *Multilayer Perceptron* (Rede Neural).
- 4 **Validação:** *K-Fold Cross-Validation* ( $k = 5$ ) com 20 execuções independentes.

# Resultados: Cenário Original (Desbalanceado)

- Treinamento realizado sem técnicas de balanceamento.
- Observa-se uma acurácia alta, porém enganosa.

**Table:** Desempenho no Conjunto de Teste - Dados Originais

<b>Classificador</b>	<b>Acurácia (Teste)</b>	<b>Recall (Validação)</b>
Random Forest	<b>81.67%</b>	$\approx 70\%$
MLP	80.00%	$\approx 69\%$

**Diagnóstico:** O modelo adota uma postura conservadora, maximizando a acurácia às custas de não identificar cerca de 30% dos óbitos.

# Melhores Hiperparâmetros (Cenário Original)

A otimização no cenário desbalanceado revelou que os modelos tendem a configurações menos complexas para evitar o *overfitting* na classe majoritária.

Table: Melhores Hiperparâmetros Encontrados (Sem Balanceamento)

Modelo	Configuração Otimizada
Random Forest	n_estimators: 29, max_depth: 6 min_samples_split: 3
MLP	activation: 'relu', hidden: (12, 15) alpha: 0.0022

- **Random Forest:** Convergiu para árvores mais rasas (*depth* 6), indicando menor necessidade de capturar fronteiras complexas.
- **MLP:** A função `relu` foi preferida, comum em dados com distribuições padrão, diferente da `tanh` exigida pelo ADASYN.

# Resultados: Cenário Balanceado (ADASYN)

- Introdução de dados sintéticos pelo ADASYN no treino.
- Inversão do comportamento do Random Forest.

**Table:** Desempenho no Conjunto de Teste - Com ADASYN

Classificador	Acurácia (Teste)	Recall (Validação)
MLP	81.67%	76.5%
<b>Random Forest</b>	78.33%	<b>86.1%</b>

**Análise:** Houve uma queda na acurácia global do Random Forest, mas um ganho substancial na capacidade de detectar pacientes em risco (Recall).



# Otimização Evolutiva: Adaptação dos Modelos

A Evolução Diferencial identificou que a alteração na distribuição dos dados exigia novas arquiteturas.

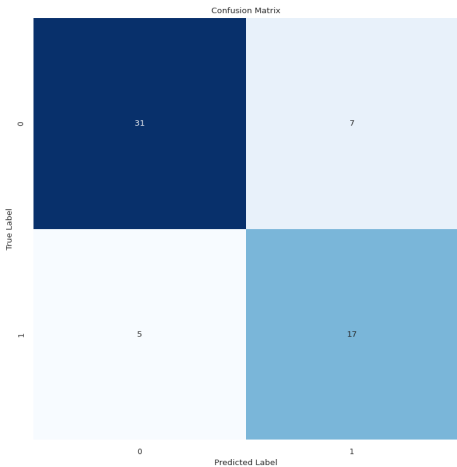
**Table:** Melhores Hiperparâmetros Encontrados (ADASYN)

<b>Modelo</b>	<b>Configuração Otimizada</b>
Random Forest	n_estimators: 29, max_depth: 9 (Aumento da complexidade da árvore)
MLP	activation: 'tanh', hidden: (11, 14) (Adaptação à não-linearidade do ADASYN)

# Análise de Erros: Matrizes de Confusão

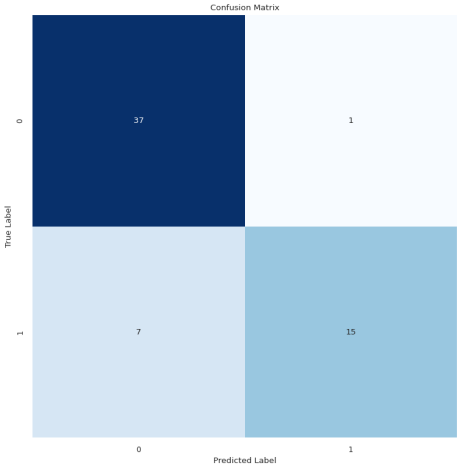
Comparação visual do impacto do balanceamento na tomada de decisão.

**Cenário Original**



Alto índice de Falsos Negativos (Risco)

**Cenário ADASYN**



Redução de Falsos Negativos (Segurança)

## Teste de Wilcoxon (Comparação de Modelos)

Avaliação da diferença de desempenho entre **Random Forest** e **MLP** no cenário balanceado (20 execuções).

- **Hipótese Nula ( $H_0$ ):** Os algoritmos têm desempenhos equivalentes.
- **Resultado ( $p$ -valor):**  $8.66 \times 10^{-5}$  ( $p < 0.05$ )
- **Conclusão:** Rejeita-se  $H_0$ . O **Random Forest** é estatisticamente superior à MLP para este problema.

## Análise de Robustez (Estabilidade)

O desvio-padrão ( $\sigma$ ) da acurácia nas múltiplas execuções indica a consistência do modelo frente à aleatoriedade da inicialização e do particionamento dos dados.

- **Desvio-padrão obtido:**  $\sigma \approx 0.01$  (1.04%)
- **Interpretação:** O valor é significativamente inferior ao limiar de 0.05, comprovando a **alta estabilidade** da otimização evolutiva.

# Impacto do Balanceamento: Comparativo Direto

## Cenário: Random Forest no Conjunto de Teste

A aplicação do ADASYN alterou o perfil de decisão do modelo, sacrificando precisão global em favor da segurança do paciente (Sensibilidade).

Métrica	Original	Com ADASYN	Impacto
Recall (Sensibilidade)	$\approx 70,0\%$	86,1%	+16,1%
Acurácia Global	81,7%	78,3%	-3,4%

Table: Comparação de desempenho (Média das execuções)

### • Interpretação Clínica:

- **Cenário Original:** Ocultava riscos. Deixava de detectar  $\approx 30\%$  dos óbitos reais.
- **Cenário ADASYN:** Prioriza a vida. O aumento drástico no Recall significa que o modelo agora identifica a grande maioria dos pacientes em risco crítico.
- *Trade-off:* A queda na acurácia deve-se ao aumento de Falsos Positivos (alarmes falsos), um custo aceitável na triagem médica.

- **Impacto do ADASYN:**

- Indispensável para o contexto médico: reequilibró as fronteiras de decisão, priorizando a detecção de riscos (redução crítica de Falsos Negativos).

- **Superioridade do Random Forest:**

- Modelo mais robusto e estável que a MLP, que exigiu mudanças estruturais complexas.
- **Resultado Validado:** Recall de 86,1% com significância estatística (Wilcoxon,  $p < 0.05$ ) e alta estabilidade (baixo desvio-padrão).
- **Justificativa:** Árvores de decisão e Ensembles (como RF) lidam melhor com dados tabulares de pequeno porte (299 registros) do que Redes Neurais profundas, que exigem grandes volumes de dados para generalizar sem *overfitting*.

- **Eficácia da Otimização:**

- A Evolução Diferencial foi crucial para encontrar configurações não triviais (e.g., árvores mais profundas) para capturar as nuances dos dados sintéticos.

# Obrigado!