

# RELATÓRIO EXECUTIVO

*[Modelo Machine Learning - PrevenCor]*



# Sumário Executivo

Este relatório apresenta os resultados de um projeto de Machine Learning focado no desenvolvimento e avaliação de um modelo de Regressão Logística para prever o risco de doenças cardíacas. Nosso objetivo principal foi construir uma ferramenta preditiva robusta e, crucialmente, interpretável, que possa auxiliar na identificação precoce de indivíduos em risco, otimizando recursos e, potencialmente, salvando vidas.

## Principais Insights e Resultados:

- **Dados Abrangentes:** O modelo foi treinado e avaliado em um conjunto de dados anonimizados de 917 pacientes, cobrindo dados clínicos essenciais de hospitais da Hungria, Suíça e Estados Unidos.
- **Fatores Chaves:** Características clínicas como Inclinação do Segmento ST, Tipo de dor no peito, Frequência cardíaca máxima) e Depressão do ST induzida por exercício demonstraram ser os preditores mais significativos.
- **Desempenho:** O modelo preditivo baseado em Regressão Logística alcançou uma Média de Precisão (AP) de 91% na Curva *Precision-Recall*, considerado um resultado robusto para problemas de classificação em saúde.
- **Capacidade de Detecção:** O modelo demonstra uma Sensibilidade (*Recall*) de 92%, indicando que é altamente eficaz em identificar a grande maioria dos casos de doença cardíaca (92% dos casos reais são corretamente detectados).
- **Interpretabilidade:** Ferramentas avançadas como SHAP revelaram que valores altos de Inclinação do Segmento ST Elevada (*UP*) e Angina Atípica (Assintomática) aumentam significativamente o risco de desenvolver doença cardíaca, enquanto Inclinação do Segmento ST Normal (*Flat*) e pacientes do sexo feminino diminuem o risco.

## Insights Gerais

Durante a etapa da Análise Exploratória dedicamos foco na compreensão profunda dos dados, identificação de padrões e insights iniciais, que se traduzem em:

**Dados Demográficos:** Avaliamos 917 registros de pacientes. Observamos um desequilíbrio de gênero, com 79% dos registros sendo de pessoas do gênero masculino e 21% do feminino. (Isso foi considerado no pré-processamento para evitar vieses).

**Idade:** A distribuição de idades é aproximadamente normal, com a mediana de 54 anos. O intervalo interquartil (IQR) de 28 a 77 anos indica que 50% dos pacientes estão nessa faixa etária, com 25% abaixo de 47, e 75% com até 60 anos.

**Sintomas Prevalentes:** 54% dos pacientes são assintomáticos, e 22% apresentam dor não anginosa. Isso significa que, em 76% dos casos, os pacientes não relataram sintomas típicos cardiovasculares, ressaltando a importância de uma ferramenta preditiva.

**Depressão do Segmento ST:** 40% dos casos apresentaram depressão nula do segmento ST, e 50% ficaram abaixo de 0.6mm (considerado normal). No entanto, 25% dos casos tiveram valores acima de 1.5mm (indicando uma diferença clinicamente relevante), com alguns outliers chegando a 6.2mm, o que reforça a relevância desta característica para a detecção de problemas cardiovasculares.

**Público:** 92 Pessoas acessam, gerando um volume de 382 eventos no mês. O público principal utiliza para monitorar a maturidade de software das equipes, exibindo indicadores criados pelo Google.

## Avaliação do Modelo de Regressão Logística

A escolha da Regressão Logística se deu por sua simplicidade, robustez e, principalmente, sua alta interpretabilidade, o que é um fator crítico em aplicações de saúde onde a compreensão das decisões do modelo é tão importante quanto a precisão das previsões. Modelos complexos ("caixas pretas") podem ser difíceis de justificar e auditar em ambientes regulamentados.

### Métricas de Desempenho:

O modelo foi avaliado em um conjunto de dados de teste (dados não vistos) e através de validação cruzada para garantir sua robustez:

Métricas de Desempenho		
Medidas	Teste	Validação Cruzada (5 dobras)
<i>Recall</i>	0.93	0.92
<i>Precision</i>	0.86	0.85
<i>F1-Score</i>	0.89	0.88
<i>Average Precision (AP)</i>	0.91	0.90

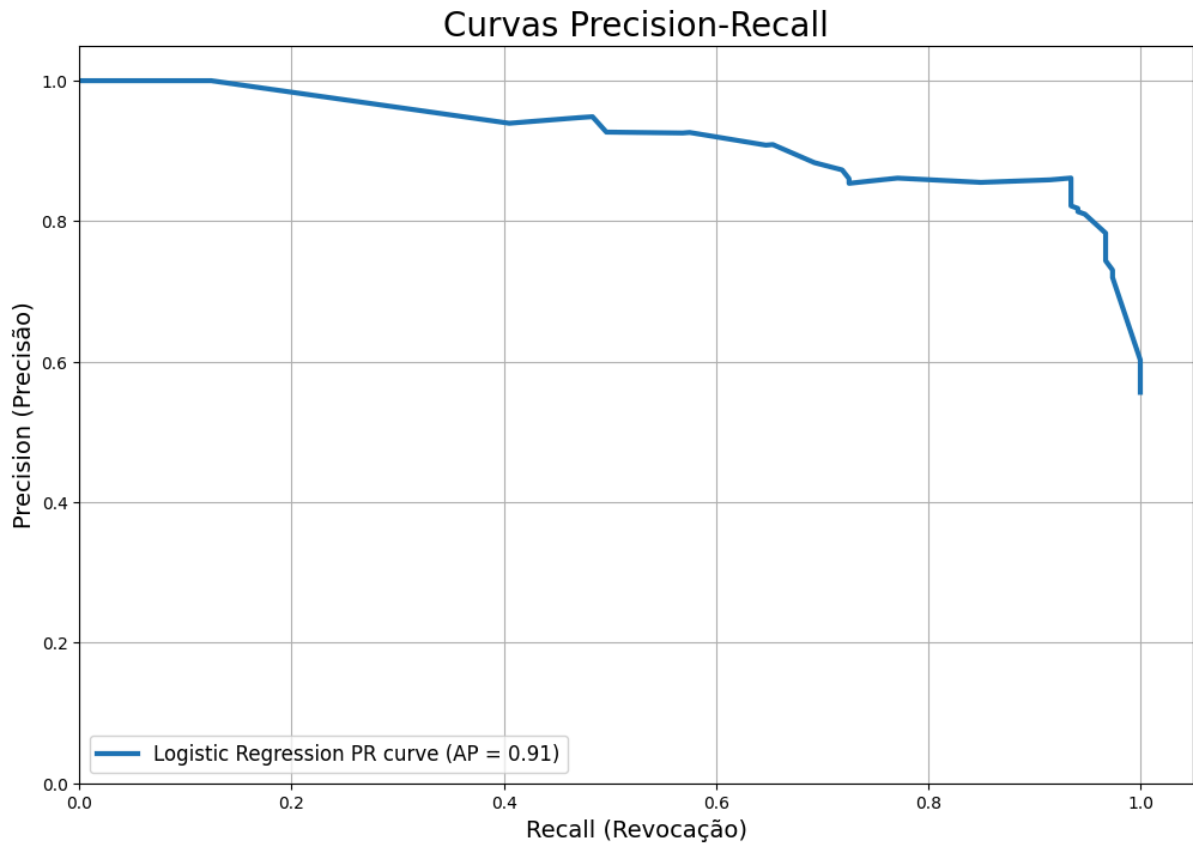
Bench no mercado sugere que um AP acima de 0.85 é considerado um ótimo resultado em problemas de classificação, especialmente para a área da saúde, indicando um modelo robusto. Ademais, a consistência entre os resultados de teste e validação cruzada confirma a estabilidade e boa generalização do modelo.

## Matriz de Confusão:

Matriz de Confusão (normalizada por linha)		
	Previsão: Não doença	Previsão: Com doença
Real: Não doença	0.81	0.19
Real: Com doença	0.065	0.93

- O modelo identificou corretamente 93% dos pacientes que realmente tinham a doença. Este é um resultado crucial para um problema de saúde, pois minimiza o risco de falsos negativos (não detectar a doença quando ela existe).
- Apenas 6.5% dos pacientes com a doença foram erroneamente classificados como saudáveis. Embora seja uma taxa baixa, em saúde, cada falso negativo é crítico e pode ser um ponto de discussão.
- 81% dos pacientes sem a doença foram corretamente identificados.
- 19% dos pacientes sem a doença foram erroneamente classificados como doentes. Este é um ponto a ser considerado, pois pode levar a exames desnecessários e ansiedade para o paciente, embora seja um "custo" frequentemente aceitável para garantir um alto *Recall* em diagnósticos.

## Curva Precision-Recall:



A curva demonstra um bom equilíbrio entre Precisão e Recall. Podemos ajustar o limiar de classificação para priorizar mais a detecção (*Recall* - Sensibilidade) ou a certeza da previsão (*Precision* - Precisão), dependendo da política de risco da instituição.

### Como por exemplo:

- É possível definir o limiar de corte quando a Precisão (*Precision*) é de 80%, elevando a Sensibilidade (*Recall*) para 94.77%.
- Ou ainda é possível definir o limiar de corte quando a Precisão (*Precision*) é de 78%, elevando a Sensibilidade (*Recall*) para 96.73%.

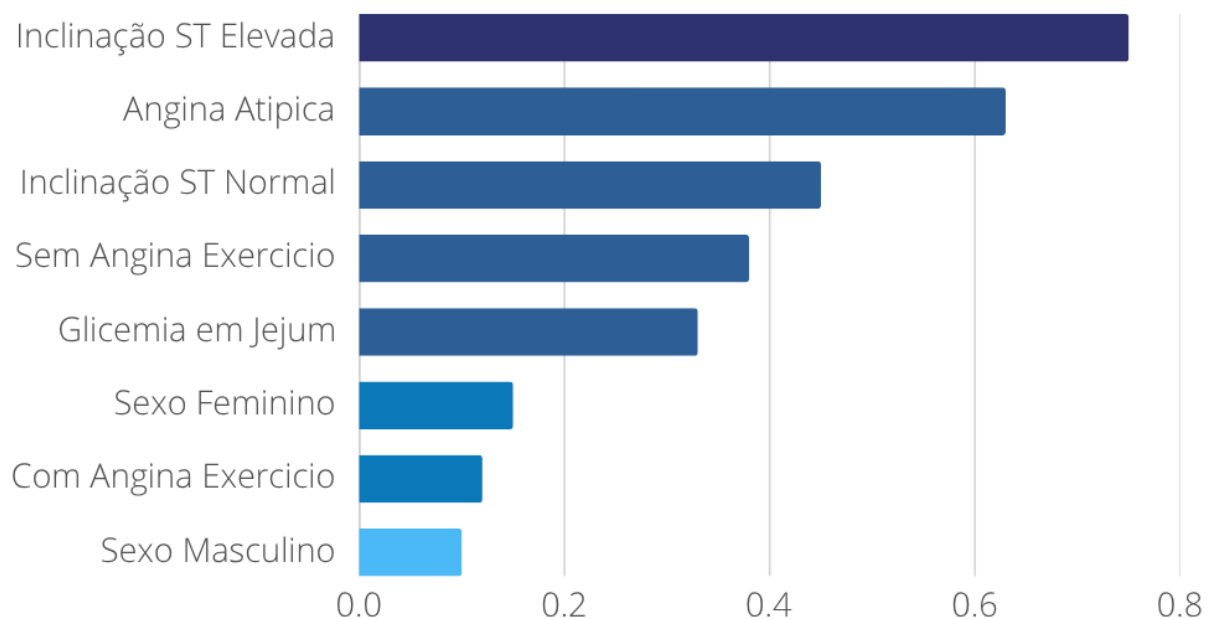
Se evitar Falsos Negativos for a prioridade máxima (e.g., para não perder um diagnóstico), o modelo pode operar em um ponto de alto Recall, aceitando um ligeiro aumento nos Falsos Positivos.

Se evitar Falsos Positivos for mais importante (e.g., para reduzir custos de exames desnecessários), o modelo pode operar em um ponto de alta precisão, aceitando um Recall ligeiramente menor.

## Interpretabilidade do Modelo:

Utilizamos SHAP (SHapley Additive exPlanations) para entender as decisões do modelo:

### Fatores mais influentes:



- **Inclinação do Segmento ST Elevado (*Up*):** Fortemente associado a uma maior probabilidade de doença cardíaca.
- **Angina Atípica (Assintomático):** Fortemente associado a uma maior probabilidade de doença cardíaca.
- **Inclinação do Segmento ST Normal(*Flat*):** Fortemente associado a uma menor probabilidade de doença cardíaca.
- **Sem Angina de Exercício:** Associado a uma menor probabilidade de doença cardíaca.
- **Glicemia em Jejum:** Valores altos aumentam a probabilidade, valores baixos diminuem.

- **Sexo Feminino:** Associado a uma menor probabilidade de doença cardíaca (em comparação com o sexo masculino).
- **Sexo Masculino:** Associado a uma maior probabilidade de doença cardíaca (em comparação com o sexo masculino).

## Oportunidades de Melhoria

O modelo atual demonstra uma performance robusta, mas há sempre espaço para aprimoramento que serão contemplados na próxima versão do modelo:

**Fortalezas:** O modelo é extremamente eficaz em identificar pacientes com doença cardíaca (Alta Sensibilidade de 0.93). Essa característica é fundamental em ambientes de diagnóstico, onde a falha em identificar um caso positivo pode ter consequências graves.

### Pontos a Considerar:

- **Falsos Positivos (19%):** Quase um quinto das pessoas sem a condição são erroneamente identificadas como doentes. Embora o "custo" disso possa ser aceitável dependendo do contexto (exames adicionais, ansiedade do paciente), futuras otimizações poderiam focar em reduzir essa taxa, talvez com um ajuste no limiar de classificação ou incorporando features adicionais.
- **Falsos Negativos (6.5%):** Apesar de ser uma taxa relativamente baixa, em saúde, cada falso negativo é crítico. Uma discussão com especialistas da área é crucial para determinar se essa taxa é aceitável ou se o modelo precisa ser ainda mais conservador para evitar FNs, mesmo que isso acarrete um leve aumento nos FPs.

### Potenciais Melhorias:

- Explorar algoritmos de Machine Learning adicionais (e.g., modelos de boosting como XGBoost ou LightGBM) que podem oferecer ganhos de desempenho.
- Realizar engenharia de features mais avançada, criando novas variáveis a partir das existentes que possam capturar relações mais complexas.