

# AnginaAlert: Identificação de Angina Com Inteligência Artificial

Guilherme A. P. Dias<sup>1</sup>, Maria Eduarda de O. Carpejani<sup>2</sup>, Robson P. Bonidia<sup>3</sup>

<sup>1</sup>Departamento de Ciência de Dados — Faculdade de Tecnologia de Ourinhos  
Av. Vitalina Marcusso, 1440, Campus Universitário — Ourinhos–SP, Brasil

**Abstract.** *Heart diseases are the leading cause of death worldwide. According to the Ministry of Health, approximately 14 million Brazilians have some cardiac condition, and about 400,000 die annually in the country due to these illnesses, which corresponds to about 30% of all deaths. Considering this, the present article introduces a tool, AnginaAlert, that can efficiently and accurately predict whether a patient is suffering from angina, using Artificial Intelligence to assist professionals in making more informed and swift decisions. This solution has the potential to lead to quicker and more precise consultations and diagnoses, relieving the pressure on healthcare facilities and enhancing the population's quality of life.*

**Resumo.** *As doenças cardíacas são as principais causas de morte em todo o mundo. De acordo com o Ministério da Saúde, aproximadamente 14 milhões de brasileiros possuem alguma condição cardíaca, e cerca de 400 mil falecem anualmente no país devido a essas enfermidades, o que corresponde a cerca de 30% do total de mortes. Considerando isso, o presente artigo apresenta uma ferramenta, o AnginaAlert, que pode prever, de forma eficiente e acurada, se um paciente está sofrendo de angina, utilizando Inteligência Artificial para auxiliar profissionais a tomar decisões mais informadas e ágeis. Esta solução tem potencial para resultar em atendimentos e diagnósticos mais rápidos e precisos, desafogando os pontos de atendimento e impulsionando a qualidade de vida da população.*

## 1. Introdução

Conforme a Sociedade Brasileira de Cardiologia (SBC) [Cardiol 2021], estima-se que ocorram cerca de 1100 mortes por dia por causas cardíacas, somando o dobro de mortes em relação aquelas em decorrência a todos os tipos de câncer juntos. Segundo um levantamento realizado pelo Instituto Nacional de Cardiologia (INC) [Cardoso 2023], entre os anos de 2017 a 2021, mais de 7 milhões de brasileiros perderam suas vidas por causa de doenças cardiovasculares. O mesmo cenário repete-se ao redor do mundo, segundo uma publicação realizada pela Organização Pan-Americana de Saúde [Organização Pan-Americana da Saúde 2020] as doenças cardiovasculares, assim como no Brasil, prevalecem como as principais causas de óbitos, sendo que cerca de 9 milhões de pessoas morreram em 2019, representando 16% do total das mortes no mundo naquele ano.

Uma possível solução para o problema em questão é o diagnóstico precoce da população, que pode ser feito com o auxílio da tecnologia, como é o caso da Inteligência

Artificial (IA) [Lobo 2017]. A IA tem ganhado cada vez mais espaço no mundo e principalmente no campo da saúde, com diversos exemplos de aplicação na medicina, como o sistema *Watson Oncology* desenvolvido pela IBM [IBM 2023], e o Projeto *Hanover* realizado pela *Microsoft* [Linn 2023]. No entanto, existem grandes desafios, entre eles: (1) Como o uso da IA para a detecção de doenças poderia alterar o cenário vivido por milhares de brasileiros? (2) O país teria a capacidade de oferecer o acesso a essas novas tecnologias para a população?

Infelizmente, o Brasil também sofre de uma doença grave e histórica, a desigualdade social. Recentes pesquisas realizadas pelo Laboratório das Desigualdades Mundiais (do inglês, *World Inequality Lab*) [Fernandes 2021], evidencia ainda mais a desigualdade vivida pelos brasileiros, como por exemplo, as regiões Sudeste e Sul sendo mais desenvolvidas e possuindo maior infraestrutura em comparação com as regiões Norte e Nordeste. Além disso, a desigualdade social é um dos principais fatores que levam à dificuldade de tratamento médico para a população, pois a parcela que mais precisa de atendimento médico é a mesma que menos possui acesso a ele [Tomasiello et al. 2023, Oliveira et al. 2023].

Considerando isso, o presente artigo, denominado **AnginaAlert**, possui como objetivo o desenvolvimento de uma nova ferramenta de identificação de condições cardíacas em pacientes, focado na Angina. Essa ferramenta utilizará dados da condição clínica dos pacientes obtidos pela triagem, exames e profissional responsável, com a finalidade de auxiliar no diagnóstico médico. Outra objetivo é desenvolver uma interface que possa ser utilizada em unidades de saúde que não possuem grande infraestrutura, impulsionando a qualidade de vida das pessoas.

## 2. Trabalhos Relacionados

Durante o processo de revisão da literatura, foram encontrados trabalhos que abordam temas e assuntos semelhantes ao desenvolvido. Alguns exemplos de trabalhos relacionados estão listados na Tabela 2. Esses trabalhos em sua maioria utilizaram IA para a identificação de alguma condição e/ou doença cardíaca ou apenas para a análise dos resultados e das estatísticas geradas pelo algoritmo utilizado. Entretanto, nenhum dos trabalhos encontrados tinha como premissa a identificação da angina em pacientes com o intuito de auxiliar o sistema de saúde a melhorar sua qualidade e agilidade no processo de diagnóstico e tratamento do paciente, democratizando o acesso da população aos benefícios da IA.

## 3. Materiais e Métodos

### 3.1. Base de Dados

A base de dados utilizada foi obtida na plataforma online Kaggle [Anand 2017], sendo publicada no ano de 2017, possuindo 303 registros de pacientes de localidades diferentes ao redor do mundo. As variáveis presentes na base de dados podem ser observadas na lista abaixo.

- Idade do paciente, Sexo do paciente, Tipo de dor no peito, Pressão arterial, Colesterol — mg/dL, Presença de açúcar no sangue em jejum superior a 120 mg/dL, Resultados eletrocardiográficos de pacientes em repouso, Frequência cardíaca alterada, Angina induzida por exercício, Depressão do segmento ST induzida por

**Tabela 1. Trabalhos Relacionados**

<b>Artigo</b>	<b>Objetivo</b>
[Shah et al. 2020]	Utiliza técnicas de mineração de dados e Aprendizado de Máquina (AM) para analisar dados médicos complexos e prever doenças cardíacas
[Ribeiro et al. 2020]	Compara a capacidade de uma Rede Neural treinada em um amplo conjunto de dados de ECG de 12 derivações para reconhecer seis tipos de anormalidades
[Ahsan and Siddique 2022]	Analisa a detecção de doenças cardíacas por meio de AM, usando dados de ECG e informações de pacientes
[Singh and Kumar 2020]	Testa a detecção de doenças cardíacas utilizando diferentes técnicas e algoritmos de AM
[Gavhane et al. 2018]	AM para identificação de doenças cardíacas utilizando as informações básicas para a identificação, como: idade, sexo, frequência cardíaca, etc.

exercício físico, Inclinação do segmento ST durante o exercício de pico, Resultado do teste de estresse cardíaco, Variável alvo (indicando a presença ou ausência de angina).

Durante a análise exploratória dos dados para avaliar a qualidade e a viabilidade da base de dados para o projeto, não foi constatada a presença de dados nulos ou faltantes, concluindo-se, que a base de dados encontrava-se completa. Para obter uma compreensão mais detalhada sobre as informações contidas na base de dados e avaliar a proporcionalidade dos dados, foram gerados gráficos individuais para cada variável. Além dos gráficos, também foi realizada uma análise estatística. Por fim, pode-se concluir que na base de dados havia variáveis que não eram proporcionais, o que poderia ocasionar no desenvolvimento de uma IA enviesada. Com o objetivo de criar uma IA responsável e imparcial, foi realizada a aplicação de técnicas para a resolução do problema, por exemplo, o uso híbrido de *undersampling* e *oversampling*. Para a aplicação dessas abordagens foram aplicados os seguintes algoritmos: *RandomOverSampler*, *RandomUnderSampler* e *FitResample*. O objetivo era criar uma base de dados mais balanceada para o desenvolvimento e treinamento da IA.

### 3.2. Avaliação Experimental

Para avaliar o desempenho da presente proposta, foram utilizados quatro algoritmos de AM: *XGBoost*, *Decision Tree*, *CatBoost* e *Random Forest*. O critério de avaliação baseia-se em métricas específicas — acurácia, precisão, *recall* e *F1-Score* — para determinar qual modelo oferece o desempenho mais eficaz durante as fases de treinamento e teste. Além disso, cada modelo foi submetido a um rigoroso processo de treinamento e teste, utilizando a base de dados proposta anteriormente.

## 4. Comparação dos Modelos

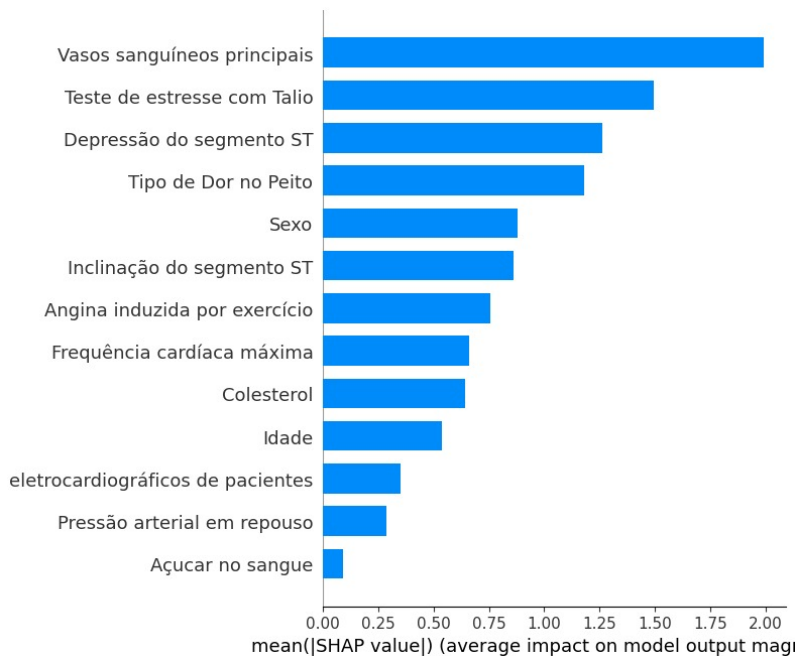
Para início dos experimentos, a base de dados foi dividida em duas partes, onde 80% foi utilizada para o treinamento dos algoritmos de AM, e os 20% restante dos dados foram

utilizados para a realização dos testes. No conjunto de treinamento, foi aplicada validação cruzada, com o objetivo de compreender como era a generalização dos algoritmos de AM. Os algoritmos que obtiveram a maior acurácia durante o seu treinamento, foram o *CatBoost* e *Random Forest*. Após a avaliação inicial do conjunto de treinamento, os algoritmos foram aplicados no conjunto de testes, conforme apresentado na Tabela 2:

**Tabela 2. Resultados dos algoritmos no conjunto de teste**

Modelo	Acurácia	Precisão	Recall	F1-Score
XGBoost	0.82	0.86	0.78	0.82
Decision Tree	0.75	0.84	0.66	0.74
CatBoost	<b>0.85</b>	<b>0.87</b>	<b>0.84</b>	<b>0.86</b>
Random Forest	0.84	0.84	0.84	0.84

Novamente, os melhores algoritmos foram o *CatBoost*, com 85% de acurácia, seguido pelo *Random Forest*, com 84%, e pelo *XGBoost* com 82%. Sendo assim, o algoritmo selecionado para o desenvolvimento da ferramenta foi o *CatBoost*, que obteve os melhores resultados tanto durante o treinamento quanto nos testes realizados. Adicionalmente, foi realizada uma análise detalhada para determinar a importância das diversas características utilizadas pelo algoritmo para tomada de decisão. O objetivo desta análise é identificar quais atributos contribuem mais significativamente para a precisão e eficácia do algoritmo, conforme apresentado na Figura 1.



**Figura 1. O gráfico apresenta o nível de importância que cada variável possui na base de dados selecionada.**

Com base no gráfico de importância de características, pode-se concluir que as variáveis, vasos sanguíneos principais, teste de estresse com Talio (avaliação de circulação sanguínea no músculo sanguíneo) e depressão do segmento ST possuem os maiores índices de relevância na base de dados.

## 5. AnginaAlert

Para a utilização do sistema de identificação foi desenvolvido um site acessível aos profissionais de saúde (Figura 2), que ainda encontra-se em testes, no qual os profissionais poderão inserir as informações necessárias coletadas durante a consulta e a triagem do paciente. Ao fornecer os dados requeridos, o algoritmo do sistema calcula a probabilidade do paciente em questão possuir ou não angina. O resultado obtido servirá como uma ferramenta auxiliar para o diagnóstico final realizado pelo profissional.

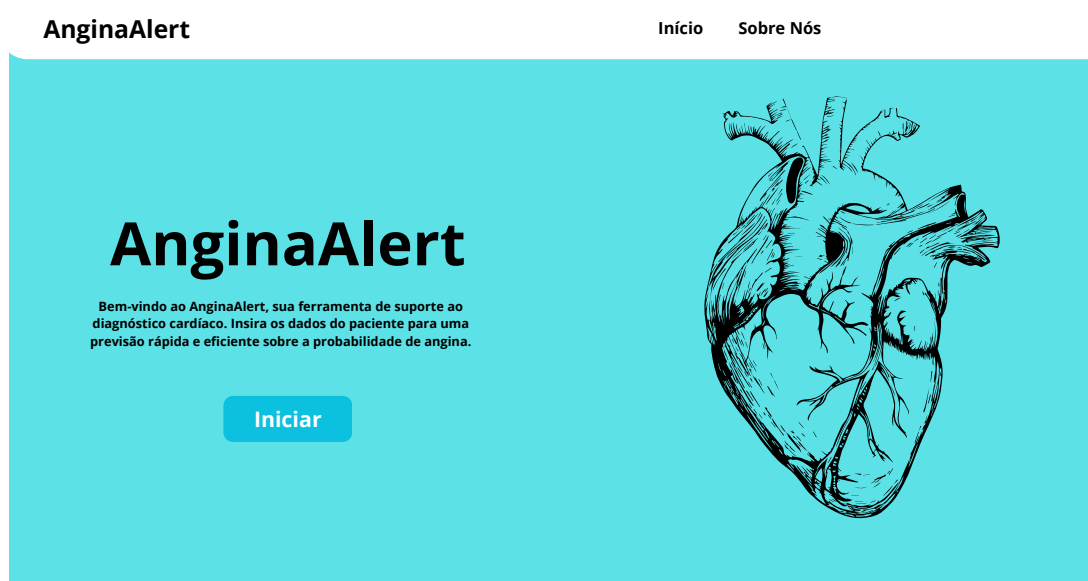


Figura 2. Tela inicial da ferramenta AnginaAlert

## 6. Conclusão

Este estudo demonstra a viabilidade de utilizar técnicas de IA, mais especificamente AM, na identificação de problemas ou condições clínicas. A identificação e diagnóstico precoce possui potencial para reduzir a incidência de óbitos causados por problemas cardíacos ou erros médicos. Além disso, ao democratizar o acesso a essa tecnologia, especialmente em populações mais carentes e afastadas de grandes centros de tratamento, pode-se induzir uma transformação substancial tanto no diagnóstico quanto no tratamento dos pacientes. Finalmente, é importante ressaltar que a ferramenta *AnginaAlert* foi desenvolvida para auxiliar no processo de diagnóstico médico; contudo, o diagnóstico final deve sempre ser realizado por um profissional qualificado da área da saúde.

## Referências

- Ahsan, M. M. and Siddique, Z. (2022). Machine learning-based heart disease diagnosis: A systematic literature review. *Artificial Intelligence in Medicine*, 128:102289.
- Anand, N. (2017). Heart attack prediction. Disponível em: <https://www.kaggle.com/datasets/imnikhilanand/heart-attack-prediction/code>. Acessado em 22 de Novembro de 2023.

- Cardiol, P. (2021). Aumenta o número de mortes por doenças cardiovasculares no primeiro semestre de 2021. Disponível em: <https://www.portal.cardiol.br/post/aumentao-n%C3%BAmero-de-mortes-por-doen%C3%A7as-cardiovasculares-no-primeiro-semestre-de-2021>. Acessado em 26 de Novembro de 2023.
- Cardoso, R. (2023). Internações por infarto aumentam mais de 150% no Brasil. Disponível em: <https://agenciabrasil.ebc.com.br/saude/noticia/2023-07/internacoes-por-infarto-aumentam-mais-de-150-no-brasil>. Acessado em 26 de Novembro de 2023.
- Fernandes, D. (2021). 4 dados que mostram por que Brasil é um dos países mais desiguais do mundo, segundo relatório - BBC News Brasil. Acessado em: 24 jan. 2022.
- Gavhane, A., Kokkula, G., Pandya, I., and Devadkar, K. (2018). Prediction of heart disease using machine learning. In *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pages 1275–1278.
- IBM (2023). IBM documentation - 5725-w51 IBM Watson for Oncology. Disponível em: <https://www.ibm.com/docs/en/announcements/watson-oncology?region=CAN>. Acessado em 26 de Novembro de 2023.
- Linn, A. (2023). How Microsoft is helping to ‘solve’ cancer. *Microsoft Story Labs*.
- Lobo, L. C. (2017). Inteligência artificial e medicina. *Revista Brasileira de Educação Médica*, 41:185–193.
- Oliveira, L. G. F., Fracolli, L. A., Araújo, S. S. d., Gryscek, A. L. d. F. P. L., Silva, G. L., Lico, F. M. d. C., Silva, L. A. d., Farias, L. G., Fogueira, J. A. L., and Silva, J. C. N. (2023). Acesso e uso equitativo dos serviços de saúde: um desafio para a promoção da universalização em saúde. *Acesso à saúde: desafios, perspectivas, soluções e oportunidades na atenção primária à saúde*.
- Organização Pan-Americana da Saúde (2020). OMS revela principais causas de morte e incapacidade em todo o mundo entre 2000 e 2019. Disponível em: <https://www.paho.org/pt/noticias/9-12-2020-oms-revela-principais-causas-morte-e-incapacidade-em-todo-mundo-entre-2000-e>. Acessado em 26 de Novembro de 2023.
- Ribeiro, A. H., Ribeiro, M. H., Paixão, G. M., Oliveira, D. M., Gomes, P. R., Canazart, J. A., Ferreira, M. P., Andersson, C. R., Macfarlane, P. W., Meira Jr, W., et al. (2020). Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nature Communications*, 11(1):1760.
- Shah, D., Patel, S., and Bharti, S. K. (2020). Heart disease prediction using machine learning techniques. *SN Computer Science*, 1:1–6.
- Singh, A. and Kumar, R. (2020). Heart disease prediction using machine learning algorithms. In *2020 International Conference on Electrical and Electronics Engineering (ICE3)*, pages 452–457.
- Tomasiello, D. B., Bazzo, J. P., Parga, J. P., Servo, L. M., and Pereira, R. H. M. (2023). Desigualdades raciais e de renda no acesso à saúde nas cidades brasileiras.