

HELMAN CAMPOS MARTINS

**Avaliação prognóstica de cardiotoxicidade por meio da ecocardiografia  
utilizando aprendizado de máquina**

São Paulo

2024

**HELMAN CAMPOS MARTINS**

**Avaliação prognóstica de cardiotoxicidade por meio da ecocardiografia  
utilizando aprendizado de máquina**

Tese apresentada à Faculdade de  
Medicina da Universidade de São Paulo e  
à Universidade Federal da Paraíba no  
âmbito do Projeto de Doutorado  
Interinstitucional - DINTER para obtenção  
do título de Doutor em Ciências.

Programa de Cardiologia

Orientadora: Profa. Dra. Vera Maria Cury  
Salemi

Coorientador: Prof. Dr. Marcelo Dantas  
Tavares de Melo

São Paulo

2024

**Dados Internacionais de Catalogação na Publicação (CIP)**

Preparada pela Biblioteca da  
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Martins, Helman Campos

Avaliação prognóstica de cardiotoxicidade por meio da ecocardiografia utilizando aprendizado de máquina / Helman Campos Martins; Vera Maria Cury Salemi, orientadora; Marcelo Dantas Tavares de Melo, coorientador. – São Paulo, 2024.

Tese (Doutorado) – Programa de Cardiologia. Faculdade de Medicina da Universidade de São Paulo; Projeto de Doutorado Interinstitucional – DINTER. Universidade Federal da Paraíba, 2024.

1. Antraciclina 2. Câncer de mama 3. Cardiotoxicidade 4. Inteligência artificial 5. Aprendizado de máquina I. Salemi, Vera Maria Cury, orient. II. Melo, Marcelo Dantas Tavares de, coorient. III. Título.

USP/FM/DBD- 460/24

Responsável: Daniela Amaral Barbosa, CRB-8 7533

Aprovado em:

Banca Examinadora

Prof. Dr.(a) \_\_\_\_\_

Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_

Profa. Dr.(a) \_\_\_\_\_

Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_

Prof.Dr.(a) \_\_\_\_\_

Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_

## **DEDICATÓRIA**

Primeiramente, gostaria de agradecer a minha amada esposa Ana Cláudia e aos meus filhos, Elizabeth Helena, Ana Caroline e Helman Fernando, pelo apoio e compreensão em todos os momentos.

A Professora Doutora Vera Maria Cury Salemi, nossa orientadora, sempre atenciosa e disponível para compartilhar o seu conhecimento.

Ao Professor Doutor Marcelo Dantas Tavares de Melo, nosso coorientador, presente em todas as etapas da nossa pesquisa.

A Professora Doutora Thaís Gaudêncio do Rêgo, nossa referência em Inteligência Artificial na Universidade Federal da Paraíba.

Aos alunos de iniciação científica, Ewerton Freitas de Medeiros e Yann Nicholas Gomes da Nóbrega jovens pesquisadores, que muito contribuíram na aplicação da Inteligência Artificial na pesquisa.

Ao Professor Doutor André Luiz Cerqueira de Almeida, por gentilmente ter cedido os dados da coorte histórica.

Aos meus amados netos Gabriel e Luiz Mateus, que trouxeram mais alegria e doçura na minha vida.

O presente trabalho faz parte de tese de doutorado apresentado à banca do Programa de Pós-Graduação em Cardiologia do Instituto do Coração da Faculdade de Medicina da Universidade de São Paulo, com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES)-código de financiamento 01.

MARTINS, Helman Campos. Avaliação prognóstica de cardiotoxicidade por meio da ecocardiografia utilizando aprendizado de máquina. [tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2024.

## RESUMO

**Introdução:** A identificação precoce de sinais subclínicos de cardiotoxicidade poderá mudar a história natural dos pacientes em tratamento de câncer de mama com antracíclicos, com medidas terapêuticas para evitar a piora da função ventricular e óbito. O papel do aprendizado de máquina tem sido pouco estudado para a otimização da ecocardiografia no diagnóstico da cardiotoxicidade. **Objetivos:** O estudo foi proposto para identificar variáveis ecocardiográficas relacionadas com cardiotoxicidade e/ou óbito, analisadas por aprendizado de máquina (*machine learning*). **Métodos:** O estudo foi composto inicialmente por uma coorte histórica de 101 pacientes portadoras de câncer de mama, no período de fevereiro de 2010 a junho de 2011, submetidas a tratamento quimioterápico com antracíclicos, sendo realizada avaliação clínica, eletrocardiograma e ecocardiograma completo, incluindo Doppler tecidual e índices de deformação miocárdica. Os exames foram realizados antes da primeira sessão, três, seis, 12 e 24 meses, sendo que, completaram o estudo 58 pacientes. **Resultados:** A idade média das pacientes foi de  $52,49 \pm 12,97$  anos. Assim, de nove pacientes com queda de dez pontos percentuais e uma FEVE menor do que 50% no último momento, somaram-se oito com redução do SLG maior que 15% entre o primeiro e subsequentes momentos do estudo, totalizando 17 pacientes agora definidos como cardiotoxicidade. Associado aos desfechos de óbito (12 pacientes) e/ou cardiotoxicidade no momento final (17 pacientes), totalizando 29 pacientes. O classificador floresta aleatória apresentou o melhor resultado do estudo com 77,78% de acurácia, 87,89% de área sob a curva ROC e 80% de *recall*, o método KNN mostrou 72,22% de acurácia, 62,31% de área sob a curva ROC e 40% de *recall*, enquanto o *XGBClassifier* apresentou 81% de acurácia, 83% de área sob a curva ROC e 66,6% de *recall*. Após a otimização as variáveis ecocardiográficas demonstradas pelos valores Shap (*SHapley Additive exPlanations*), por meio da biblioteca para explicar as previsões de modelos de aprendizado de máquina, foi observado que, além da redução da fração de ejeção de ventrículo esquerdo, índices como a excursão sistólica do plano anular da valva tricúspide (TAPSE), velocidade da onda e' da região lateral da valva mitral, o índice de desempenho miocárdico e a excursão sistólica do plano anular da valva mitral (MAPSE), apresentaram relação com cardiotoxicidade e/ou óbito. **Conclusão:** Neste estudo foi demonstrado que parâmetros ecocardiográficos de fácil obtenção pela ecocardiografia transtorácica são preditores de cardiotoxicidade e óbito. Novos estudos com amostra maior de pacientes poderão confirmar esses achados.

**Palavras-chave:** Antraciclina, Câncer de mama, Cardiotoxicidade, Inteligência artificial, Aprendizado de máquina

MARTINS, Helman Campos. Prognostic evaluation of cardiotoxicity through echocardiography using machine learning. [thesis]. São Paulo: School of Medicine, University of São Paulo;2024.

## ABSTRACT

**Introduction:** Early identification of subclinical signs of cardiotoxicity may change the natural history of patients undergoing breast cancer treatment with anthracyclines, with therapeutic measures to prevent worsening of ventricular function and death. The role of machine learning for the optimization of echocardiography in the diagnosis of cardiotoxicity has been little studied. **Objective:** The study was proposed to identify echocardiographic variables related to cardiotoxicity and/or death, analyzed by machine learning. **Methods:** The study was initially composed of a historical cohort of 101 patients with breast cancer, from February 2010 to June 2011, undergoing chemotherapy treatment with anthracyclines, with clinical evaluation, electrocardiogram and complete echocardiogram, including tissue Doppler and myocardial deformation indices. The exams were performed before the first session, three, six, 12 and 24 months, and 58 patients completed the study. **Results:** The mean age of the patients was  $52.49 \pm 12.97$  years. Thus, of the nine patients with a ten percentage point drop and a LVEF lower than 50% at the last moment, there were eight with a reduction in GLS greater than 15% between the first and subsequent moments of the study, totaling 17 patients now defined as having cardiotoxicity. Associated with the outcomes of death (12 patients) and/or cardiotoxicity at the final moment (17 patients), totaling 29 patients. The random forest classifier presented the best result of the study with 77.78% accuracy, 87.89% area under the ROC curve and 80% recall, the KNN method showed 72.22% accuracy, 62.31% area under the ROC curve and 40% recall, while the XGBClassifier presented 81% accuracy, 83% area under the ROC curve and 66.6% recall. After optimizing the echocardiographic variables demonstrated by the Shap (SHapley Additive exPlanations) values, through the library to explain the predictions of machine learning models, it was observed that, in addition to the reduction in the left ventricular ejection fraction, indices such as the tricuspid valve annular plane systolic excursion (TAPSE), e' wave velocity of the lateral region of the mitral valve, the myocardial performance index and the mitral valve annular plane systolic excursion (MAPSE), were related to cardiotoxicity and/or death. **Conclusion:** This study demonstrated that echocardiographic parameters easily obtained by transthoracic echocardiography are predictors of cardiotoxicity and death. New studies with a larger sample of patients may confirm these findings.

**Keywords:** Anthracycline, Breast cancer, Cardiotoxicity , Artificial intelligence · Machine learning.

## LISTA DE FIGURAS

<b>Figura 1:</b> Modelo de árvore de decisão .....	20
<b>Figura 2:</b> Relação gráfica entre Entropia e Critério de Gini.....	21
<b>Figura 3:</b> Representação de classificador por floresta aleatória.....	22
<b>Figura 4:</b> Representação gráfica do método KNN.....	23
<b>Figura5:</b> Fluxograma com as etapas de pré-processamento dos dados.....	36
<b>Figura 6:</b> Matriz de confusão obtida pelo K Neighbors classifier .....	43
<b>Figura 7:</b> Curva ROC AUC para o modelo KNN.....	44
<b>Figura8:</b> Matriz de confusão para o modelo floresta aleatória.....	47
<b>Figura 9:</b> Curva ROC para o modelo <i>floresta aleatória</i> .....	48
<b>Figura 10:</b> Comparação da Acurácia, AUC ROC, Recall e Precisão dos modelos de predição, em porcentagem.....	48
<b>Figura 11:</b> Gráfico da média absoluta de importância de cada variável pelo modelo SHAP para a floresta aleatória.....	50
<b>Figura 12:</b> <i>Summary plot</i> no modelo SHAP para análise de importância de variáveis a partir do Random Forest.....	51

## LISTA DE ABREVIATURAS

- a'**: Pico de velocidade da onda a' obtida pelo Doppler tecidual
- A**: Pico de velocidade da onda A do fluxo mitral
- AE**: Átrio esquerdo
- Ao**: Raiz da aorta
- CIR\_ABD**: Circunferência abdominal
- DdVE**: Diâmetro diastólico final do ventrículo esquerdo
- DsVE**: Diâmetro sistólico final do ventrículo esquerdo
- DOSE\_TOTALM<sup>2</sup>**: Dose total de antraciclina utilizada por paciente
- DNA**: Ácido desoxirribonucleico
- DT**: *Decision Tree (árvore de decisão)*
- ECG**: Eletrocardiograma
- E**: Pico de velocidade da onda E do fluxo mitral
- e'**: Pico de velocidade da onda e' obtida pelo Doppler tecidual
- E/e'**: Valor médio da relação E/e' (utilizando o valor médio da onda E' nas regiões da valva mitral (septal, lateral, anterior e inferior))
- FC**: Frequência cardíaca
- FEVE**: Fração de ejeção ventrículo esquerdo
- HDPA-UNACON**: Hospital Dom Pedro de Alcantara- Unidade de Alta Complexidade em Oncologia
- IA**: Inteligência artificial
- KNN**: *K-Nearest Neighbors*
- MAPSE**: Excursão sistólica do plano anular mitral
- NLP**: *Natural language processing* (processamento de linguagem natural)
- PCA**: *Principal component analysis* (análise do componente principal)
- SLG**: *Strain* longitudinal global
- TAPSE**: Excursão sistólica de plano anular tricúspide

## **SUMÁRIO**

<b>1 INTRODUÇÃO</b>	<b>12</b>
1.1 Cardio-oncologia	12
1.2 Cardiotoxicidade	12
1.3 Ecocardiograma	12
1.4 Aprendizado de máquina	13
<b>2 OBJETIVO</b>	<b>23</b>
<b>3 METODOLOGIA</b>	<b>24</b>
3.1 Desenho do estudo	24
3.2 Material	24
3.3 Casuística	25
3.4 Critérios de inclusão	25
3.5 Critérios de exclusão	25
3.6 Exame ecocardiográfico	26
3.7 Aspectos éticos	27
3.8 Aprendizado de máquina	31
<b>4 RESULTADOS</b>	<b>43</b>
4.1 Predição por Classificadores	43
4.2 Análise das Variáveis	43
<b>5 DISCUSSÃO</b>	<b>50</b>
5.1 Comparação entre os modelos preditivos	51
5.2 Discussões acerca da análise de variáveis	53
<b>6 CONCLUSÃO</b>	<b>55</b>
<b>8 BIBLIOGRAFIA</b>	<b>56</b>

## **INTRODUÇÃO**

O significativo desenvolvimento da terapia oncológica aumentou o número de pacientes de pacientes sobreviventes acometidos por neoplasias. Os efeitos tóxicos de quimioterápicos, porém, podem deixar sequelas cardiovasculares, muitas delas irreversíveis.

### **1.1 Cardio-oncologia**

A cardio-oncologia é uma nova fronteira da cardiologia que busca identificar precocemente, diagnosticar e tratar as complicações causadas pelas terapias contra o câncer por meio de vários mecanismos, promovendo a interação entre as diversas especialidades envolvidas, estratificando o risco-benefício, uniformizando terminologias e conceitos, estimulando a pesquisa e difusão dos conhecimentos. (1)  
(2)

A implementação de medidas capazes de estratificar o risco cardiovascular em pacientes que irão submeter-se a terapia potencialmente cardiotoxica, com a colaboração multidisciplinar de oncologistas, hematologistas e cardiologistas, poderá prevenir as complicações cardiovasculares, reduzindo a necessidade de interrupção do tratamento oncológico e aumentando a sobrevida dos pacientes. (3)

### **1.2 Cardiotoxicidade**

Cardiotoxicidade é um efeito colateral importante da quimioterapia na terapêutica do câncer. Dentre as opções farmacológicas, destaca-se os antracíclicos, utilizados no câncer de mama, de útero, gástrico, linfoma e melanoma.

As antraciclina (doxorubicina, epirubicina, epirrubicina e idarubicina) estão relacionadas à disfunção ventricular irreversível em até 26% dos casos, a depender da dose aplicada. O mecanismo desse efeito deletério tem sido compreendido por meio de muitas vias, como estresse oxidativo resultando em um desequilíbrio entre as espécies nitrogenadas reativas e espécies oxigenadas reativas, inibição da topoisomerase 2, ativação de citocinas e marcadores inflamatórios como interleucinas, fator necrótico tumoral e fator nuclear k-B, que favorecem a apoptose.

Outros mecanismos descritos são a diferenciação de fibroblastos cardíacos para um fenótipo pró-fibrótico, miofibroblasto e dano a mitose do DNA, reduzindo a capacidade de diferenciação das células tronco mesenquimais, entre outros mecanismos. (4)

A recente diretriz europeia de cardio-oncologia conceituou terapia para o câncer associada a disfunção miocárdica de acordo com o quadro clínico apresentado: sintomáticos muito severos, que necessitam de suporte circulatório ou transplante cardíaco; graves, que necessitam de internação hospitalar devido a insuficiência cardíaca; moderados, com tratamento ambulatorial otimizado; e a forma leve, que não necessita de otimização terapêutica.

No grupo de assintomático foi descrito o grupo com disfunção ventricular grave, caracterizado por fração de ejeção do ventrículo esquerdo (FEVE) menor que 40%, no grupo de disfunção moderada com uma queda de 10 ou mais pontos, com uma FEVE entre 40 e 49% ou redução menor do que 10 pontos associada a um *strain* longitudinal global (SLG) maior do que 15 pontos e/ou elevação dos marcadores de necrose. O grupo de comprometimento leve apresenta FEVE maior do que 50% e elevação dos marcadores de necrose. (5)

O grande desafio do tratamento oncológico, especialmente, quando está indicado os antracíclicos, seria a prevenção dos efeitos cardiotoxicos e a sua identificação precoce, como a mudança no estilo de vida, com impacto na redução do peso, cessação do tabagismo e redução do consumo de álcool. (6)

A modulação neuro-hormonal com bloqueadores do sistema de renina-angiotensina-aldosterona, antagonistas do receptor de mineralocorticoide e betabloqueadores se mostraram eficazes em prevenir a redução da FEVE. (7)

Além disso, o uso de estratégias de redução de dose e velocidade de infusão dos antracíclicos (8), o uso de drogas protetoras como o dexrazoxane (9) e a utilização de antracíclicos lipossomais, com a mesma eficiência terapêutica e menor toxicidade em relação aos antracíclicos convencionais tem se mostrado úteis na prevenção da cardiotoxicidade. (10)

### **1.3 Ecocardiograma**

O ecocardiograma é um importante método no diagnóstico e acompanhamento dos pacientes submetidos a terapêutica oncológica, especialmente quando expostos

a agentes quimioterápicos quem podem causar cardiotoxicidade. (11)

A medida da FEVE pelo método de Simpson biplano modificado tornou-se a referência para o monitoramento e diagnóstico da cardiotoxicidade. Vale ressaltar a limitação do método pela variabilidade inter-observadores, que pode chegar em 10% e as alterações de pré e pós-carga durante o tratamento oncológico, podem afetar o resultado obtido. (12)

Destaca-se as informações obtidas pela deformação miocárdica, especialmente pelo uso SLG, obtido pela técnica de rastreamento de pontos (*speckle tracking*), apresenta capacidade de identificar disfunção subclínica do ventrículo esquerdo (13), fundamental para a prevenção e ajustes terapêuticos, que pode evitar efeitos cardiotoxícos irreversíveis. (14) (15). Os intervalos de confiança mais estreitos dessa técnica em comparação a FEVE, possibilitam que as medidas protetivas sejam implementadas mais precocemente. (16)

O estudo randomizado multicêntrico internacional SUCCOUR (*Strain Surveillance of Chemotherapy for improving Cardiovascular Outcomes*) demonstrou que a terapia quimioterápica por antracíclicos guiada por medidas do SLG, com queda superior a 12%, foi capaz de evitar redução da FEVE e cardiotoxicidade em um ano. Na evolução de três anos deste estudo, verificou-se a melhora da função ventricular em relação ao dados obtidos em um ano, mas sem diferença em relação a terapia de cardioproteção guiada pela FEVE ou pelo SLG.

#### **1.4 Aprendizado de máquina.**

A inteligência artificial (IA) é obtida a partir da composição de três elementos fundamentais, sendo eles: estatística, computação e área específica de pesquisa. Neste contexto, diversas áreas de pesquisa estão sendo diretamente impactadas pelo crescimento das aplicações de IA. (17)

Dentre tais áreas, a medicina ganha destaque ao se utilizar de aplicações voltadas para predições de eventos ou na classificação de determinados padrões comuns em certas populações, tudo isso auxiliado pelos grandes avanços alcançados na capacidade de processamento computacional atual. Logo, diagnósticos cada vez mais precisos a partir de reconhecimento automático de exames de imagem ou até mesmo, previsões de eventos patológicos com base em características específicas

de populações são algumas das aplicações cada vez mais direcionadas ao uso da IA.

A IA na saúde apresenta como aplicação potencial a pouca necessidade de se estabelecer longos códigos de programação relacionados a protocolos de saúde bastante complexos. Assim, tal ferramenta se torna útil ao fazer com que a máquina utilize conceitos matemáticos para auxiliar profissionais na tomada de decisões cada vez mais assertivas baseado no que os modelos conseguiram identificar. Assim, surge uma técnica oriunda deste campo que dá a capacidade a máquina de realizar tipos de aprendizado a partir de bases de dados previamente elaboradas, e a isso dá-se o nome de aprendizado de máquina (*machine learning*). O aprendizado de máquina pode ser dividido segundo as seguintes fases:

Fase 1- Fase de Formação: Esta é a fase em que os dados de formação são utilizados para treinar o modelo, emparelhando a entrada dada com a saída esperada. O resultado desta fase é o próprio modelo de aprendizagem;

Fase 2- Validação e Fase de Teste: Esta fase destina-se a medir a qualidade do modelo de aprendizagem que foi treinado e estima as propriedades do modelo, tais como medidas de erro, recolha, precisão e outras. Esta fase utiliza um conjunto de dados de validação, e o resultado é um modelo de aprendizagem sofisticado;

Fase 3- Fase de aplicação: Nesta fase, o modelo está sujeito aos dados do mundo real para os quais os resultados precisam ser derivados;

Para uma melhor compreensão do que seria o aprendizado de máquina, alguns conceitos devem ficar bem estabelecidos, sendo estes observados da seguinte forma simplificada:

#### I - Tipos de dados

- Estruturados: São dados encontrados de forma organizada em bancos de dados, muitas vezes no formato de tabelas, separados por linhas e colunas;
- Não estruturados: Todo o conjunto de dados que não são contidos dentro de tabelas, como imagens e áudios;

#### II - Tipos de Modelos

- Modelos lógicos: São modelos baseados em algoritmos puros que visam ajudar a derivar uma série de regras entre os dados de forma sequencial. Um exemplo deste modelo é a árvore de decisão;
- Modelos geométricos: Utilizam conceitos geométricos como linhas, planos, ou distâncias com os mais variados objetivos, sendo utilizados geralmente em grandes volumes de dados. Um exemplo é o *support vector machine* (SVM) que utiliza de abstrações geométricas a partir de planos de separação espaciais a fim de realizar a classificação de dados de acordo com certos parâmetros comuns aos mesmos;
- Modelos probabilísticos: São modelos estatísticos que empregam técnicas estatísticas baseados numa estratégia que define a relação entre duas variáveis.

### III - Tipos de problemas de aprendizagem:

- Classificação: É uma forma de identificar grupos existentes dentro de uma base de dados a partir do atributo de saída comum existente entre tais grupos. Logo, é possível identificar padrões de comportamento. Um exemplo clássico é o utilizado por bancos para verificar se o cliente com determinadas características merece ou não receber um empréstimo, podendo classificá-lo como de baixo, médio ou alto risco de inadimplência;
- Clusterização: Corresponde ao método de separar elementos em grupos semelhantes a classificação, entretanto sem depender de se ter um atributo de saída comum. Um dos métodos mais utilizados para isso é o *K-means clustering*, um exemplo clássico é o de separação de filmes por tipo de escolha de um cliente. Neste caso, não existe um rótulo específico, pois o cliente pode gostar de diferentes tipos de filmes, porém, ele e outros clientes podem ter gostos semelhantes, o que pode fazer com que fiquem separados em grupos com interesse próximo;
- Previsão: Visa determinar como certos eventos vão acontecer no futuro baseados em dados correspondentes a eventos anteriores. Todos os resultados da previsão apresentam certos graus de incerteza ou probabilidade;
- Simulação: Trata-se de uma aplicação onde existem grandes quantidades de

incerteza. A simulação corresponde a uma metodologia de solução de problemas que possui uma quantidade vasta de alternativas;

- Otimização: São técnicas que visam fazer com que determinado processo tenha os melhores resultados possíveis baseados em dados, tudo isso a partir de parâmetros matemáticos que buscam encontrar o menor erro dentro de certos valores;

#### IV - Tipos de subáreas

- Aprendizagem supervisionada: Está relacionada com a área em que se opera com dados etiquetados. Todos os dados já estão rotulados em grupos desde o início, assim o método de aprendizagem de classificação é um tipo de aprendizagem supervisionada, em que já se sabe que aquele conjunto de dados tem um rótulo em comum. Logo, o aprendizado supervisionado foca em trabalhar os dados para fazer uma ligação entre o que vem de entrada e quais os valores que estão em saída. Além disso, necessita de um operador para determinar as possíveis relações;
- Aprendizagem não supervisionada: É o tipo de aprendizado em que os dados não possuem rótulos logo de início (ex: bons pagadores vs. mal pagadores). Aqui, todos os dados vão sendo separados conforme características próprias durante a execução dos algoritmos, logo não precisam que se tenha uma designação da saída. Um exemplo é o modelo de clusterização falado anteriormente;
- Aprendizagem semi-supervisionada: Consiste na utilização de dados rotulados e não rotulados para realizar a aprendizagem. A aprendizagem semi-supervisionada obtém a sua motivação da forma humana de aprendizagem;
- Aprendizagem por reforço: Trata-se do método de aprendizagem que visa maximizar os resultados a partir de recompensas e punições por ações tomadas. Desta forma, a estratégia de aprendizagem aprende a desenvolver passos para se obter as maiores recompensas possíveis com as menores punições possíveis. Um exemplo clássico é o de robôs saindo de um ponto A para um ponto B, em que existem diversas possibilidades e tais possibilidades são dados pontos que podem diminuir ou aumentar a nota final de escolha,

assim treina-se tal equipamento para escolher o caminho com a melhor nota após experimentar diversas possibilidades existentes;

- Aprendizagem profunda ou *deep learning*: Está relacionada com as redes neurais artificiais. Este tipo de modelo computacional se transforma através de diversas técnicas, muitas delas envolvendo pesos. Assim, temos a ocorrência de IA mais adequada, visto que tal aplicação aprende com os dados e não só implementa conceitos matemáticos;

### 1.4.1 Aprendizado Supervisionado

O aprendizado supervisionado consiste no objetivo de realizar um aprendizado de máquina de forma a ter um mapeamento dos dados de entrada até uma saída de forma que os valores corretos são determinados por um supervisor. Assim, o aprendizado se dá a partir do ajuste de parâmetros para que os modelos consigam fazer previsões. (18)

Uma forma de compreender as diferenças entre os tipos de aprendizado no aprendizado de máquina, pode ser observado na Tabela abaixo:

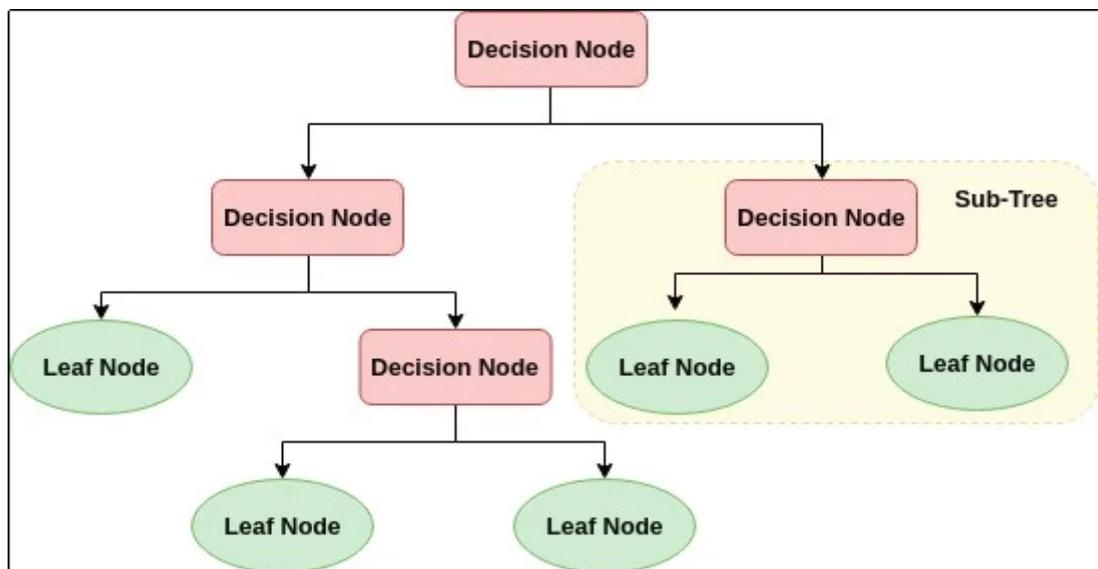
**Tabela 1** - Tipos de aplicações no aprendizado de máquina separados por principais sub-áreas

Aprendizado de Máquina		
Supervisionada	Não supervisionada	Por Reforço
Classificação Ex: <i>Random Forest</i>	Associação Ex: Apriori	Ex: <i>Q-Learning</i>
Regressão Ex: Regressão linear	Agrupamento Ex: <i>K-Means Clustering</i>	
	Detecção de desvios Ex: PCA	
	Sumarização Ex: NLP	

### 1.4.2 Aprendizado por Árvore de Decisão (*Decision Tree*)

As árvores de decisão são conhecidas como uma das técnicas mais importantes utilizadas no aprendizado de máquina. São úteis tanto na predição como também na classificação de instâncias através da representação de uma estrutura de árvore a partir da raiz até uma folha. Existem duas estruturas importantes de uma árvore de decisão - um nó e um arco que liga os nós. Cada caminho desde a raiz até uma folha corresponde a uma conjunção de condições de teste nos nós de decisão sobre o caminho e tal caminho pode ser escrito como uma regra de "se então". Para tomar uma decisão, o fluxo começa nos nós de raiz, navega para os arcos até chegar a um nó de folha, que é onde cada decisão é tomada pelo modelo. Cada nó da árvore denota o teste de um atributo, e o ramos denotam os possíveis valores que o atributo pode assumir, como demonstrado na Figura 1, vejamos:

**Figura 1** - Modelo de árvore de decisão



Legenda: Decision Node=nó de decisão, leaf node=nó de folha, sub-tree=ramo da árvore.

Para a construção de árvores de decisão torna-se necessário selecionar determinados atributos que visam saber quando se deve dividir a árvore entre os nós, sendo esta uma das características mais importantes na criação. Dentre tais atributos pode-se citar os seguintes: *Entropy* ou Entropia; *Gini Index* ou critério de Gini.

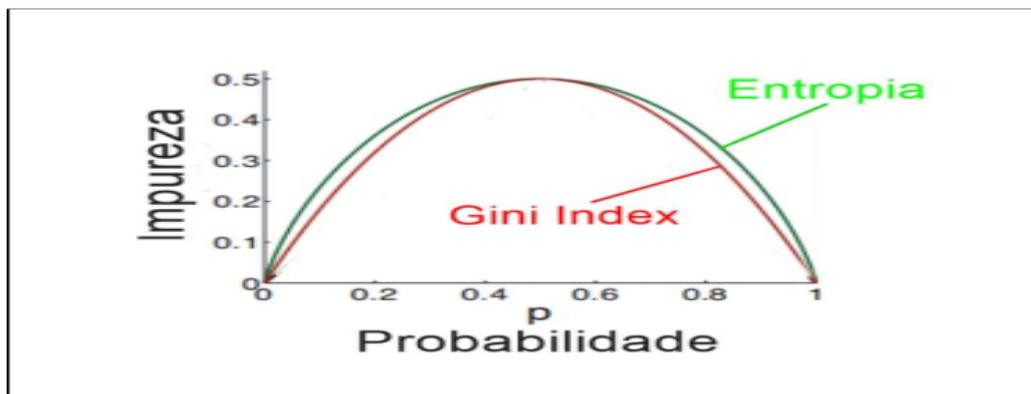
O critério denominado de entropia é usado em um algoritmo denominado de

C4.5 e mede a incerteza nos dados. Para isso, relata que o objetivo é calcular matematicamente qual evento tem a menor incerteza antes e após a divisão da árvore. Assim, a entropia se baseia na probabilidade da ocorrência de certos eventos de modo a quantificar a incerteza a partir da probabilidade de suas ocorrências. Logo, eventos com chances demasiadamente baixas ou altas de ocorrer apresentam pouca certeza, logo baixa entropia, porém eventos com probabilidade mais comum de ocorrerem apresentam mais certeza, logo, maior entropia. Abaixo pode-se observar a equação que corresponde a entropia, em que  $p$  representa a probabilidade de os eventos ocorrerem:

$$E = - \sum_{i=1}^m p_i \log_2(p_i)$$

A Figura 2 mostra o critério de impureza de Gini, que foi criado pelo estatístico italiano Corrado Gini e está relacionado com a probabilidade de dois itens ocorrerem dentro da mesma classe. Matematicamente se assemelha a entropia ao serem ambas funções convexas e muito utilizadas na prática:

**Figura 2** - Relação gráfica entre Entropia e Critério de Gini



**Fonte:** Adaptado de MEHRYAR MOHRI; AFSHIN ROSTAMIZADEH; AMEET TALWALKAR.<sup>18</sup>

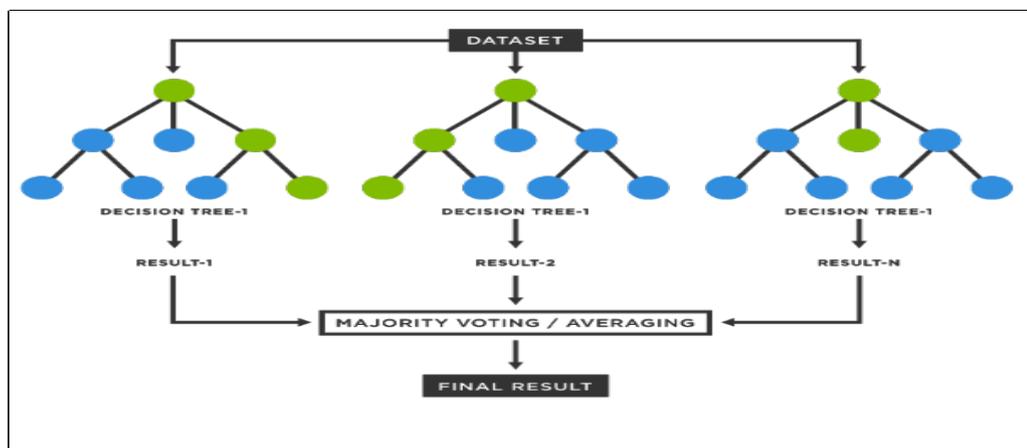
Apesar do grande uso das árvores de decisão, hoje já existem métodos que conseguiram se utilizar da capacidade de processamento de novos computadores e conseguem ampliar mais possibilidades, como acontece com o classificador floresta aleatória.

### 1.4.3 Floresta aleatória

Existem diferentes modelos de árvores de decisão e algoritmos de aprendizagem, dependendo do teste de divisão utilizado no nós de decisão e a interpolação feita nas folhas, todavia uma abordagem muito popular hoje em dia é a floresta aleatória, demonstrada na Figura 3 onde treinamos muitas árvores de decisão em subconjuntos aleatórios em que se combina as previsões com aumento significativo da acurácia nos resultados. (19)

O processo de criação ocorre quando os dados originais têm muitas dimensões, para isso, escolhe-se aleatoriamente um pequeno subconjunto das dimensões (colunas) e se constrói uma árvore. Deixa-se crescer toda a caminho sem realizar a função de poda. Por fim, faz-se a interação deste processo com a construção de centenas de árvores com um conjunto diferente de atributos de cada vez, que para o caso de predição, a média de votos de todas as árvores é reportada como o resultado da floresta aleatória.

**Figura 3** - Representação de classificador por floresta aleatória:



Legenda: *averaging*= média; *dataset*= conjunto de dados; decision tree= árvore de decisão; *final result*= resultado final; *majority voting*= maioria dos votantes (22)

### 1.4.4 K-Nearest Neighbors

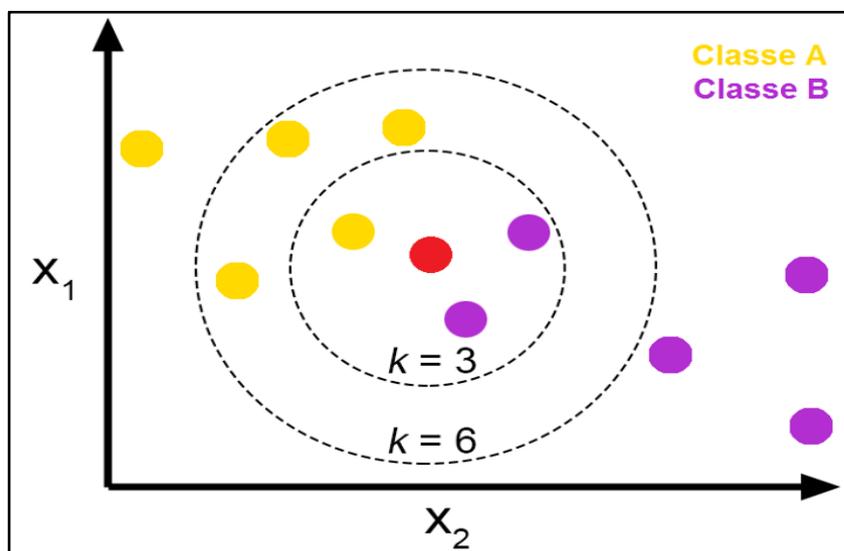
Os *K-Neighbors* (KNN) são modelos em que se interpretam instâncias em termos da formação mais semelhante de casos ao redor. Para isso, é utilizada uma

hipótese básica de que entradas com características semelhantes, têm resultados semelhantes. São também chamados métodos baseados em exemplos, memórias ou casos.

Entretanto, vale destacar que tal metodologia apresenta o uso do conceito geométrico de distância euclidiana, de forma que ele classifica um item em um determinado grupo, à medida que tal item esteja mais próximo de um conjunto de elementos com características semelhantes.

Ademais, o valor atribuído ao parâmetro K é de fundamental importância para o modelo, visto que funciona como um tipo de critério de seleção geométrica para colocar um ponto específico dentro de uma classificação de elementos mais próximos que o mesmo. Assim, sucessivos valores de K são colocados para aumentar a distância de um novo elemento a seus elementos vizinhos e, com isso, a partir de uma distância adequada, o elemento será incluído no grupo com maior quantidade de dados semelhantes. Tal representação pode ser observada na Figura 4 abaixo:

**Figura 4** - Representação gráfica do método KNN.



Logo pode-se iniciar a interpretação sobre o valor de K mais adequado para o modelo estudado, portanto, se o valor de k é muito grande, o algoritmo KNN utiliza o valor anterior e, o que pode resultar em inexatidões. Do contrário, caso o valor K seja muito pequeno, o modelo se tornaria bastante sensível a aberrantes como pouca exatidão na classificação. Daí, um valor K mais adequado tende a se situar a meio caminho do menor e maior valor. Para isso, utiliza-se técnicas (Ex.: Método do

Cotovelo) para escolher um valor neste intervalo a partir da medição de erros nos dados de formação e, assim, determinar um  $K$  com o melhor desempenho de generalização.

#### 1.4.5 XGB CLASSIFIER

O *XGBoost* (*eXtreme Gradient Boosting*) é um algoritmo de aprendizado de máquina que se destaca pela sua eficiência e eficácia em problemas de modelagem preditiva. Desenvolvido inicialmente como um projeto de pesquisa na Universidade de Washington por Tianqi Chen e Carlos Guestrin, o *XGBoost* é uma implementação otimizada do método de *boosting* em *gradient descent*. Este método é um tipo de aprendizado *ensemble* que constrói um modelo forte a partir da combinação de vários modelos simples, neste caso, árvores de decisão. O algoritmo melhora a precisão das previsões ao corrigir os erros dos modelos anteriores através de um processo iterativo e sequencial. Além disso, o *XGBoost* incorpora penalidades de regularização que ajudam a evitar o *overfitting*, permitindo que o modelo generalize melhor para dados não vistos anteriormente.

O *gradient descent* é um algoritmo de otimização iterativo que busca encontrar o mínimo de uma função. No contexto do aprendizado de máquina essa função geralmente representa o erro entre as previsões de um modelo e os valores reais. A ideia central é ajustar os parâmetros do modelo na direção oposta ao gradiente da função de erro, a fim de minimizar esse erro.

Este método ganhou espaço acadêmico e comercial por sua capacidade de lidar com dados, onde supera frequentemente outros algoritmos, e pela flexibilidade.

(20)

## **2 OBJETIVO**

Este estudo tem por objetivo identificar variáveis ecocardiográficas, em pacientes com câncer de mama, submetidas a tratamento com antracíclicos, relacionadas com cardiotoxicidade e/ou óbito, analisadas por aprendizado de máquina.

### 3 METODOLOGIA

#### 3.1 Desenho do estudo

Trata-se de um estudo de coorte histórica onde são analisados os dados ecocardiográficos e de desfechos clínicos em pacientes com câncer de mama, que foram submetidas à quimioterapia com antracíclicos. Estes dados foram obtidos em estudo conduzido pelo pesquisador Prof. Dr. André Luiz Cerqueira de Almeida, da Universidade Estadual de Feira de Santana (UEFS), sendo que reanalisamos as imagens e parâmetros ecocardiográficos dos pacientes do estudo piloto. O estudo piloto teve aprovação no comitê de ética da UEFS sob número 69195, CAEE 02590212.0.0000.0053 e o pesquisador nos forneceu carta de anuência para acesso aos dados. Cada paciente realizou uma avaliação clínica, assim como um eletrocardiograma e um ecocardiograma antes e após a primeira sessão de quimioterapia, assim como após três, seis, 12 e 24 meses do início do tratamento.

Foi investigado o valor preditivo das variáveis ecocardiográficas, tendo como desfecho mortalidade total, morte por causa cardiovascular e redução da fração de ejeção do ventrículo esquerdo maior que 10% do basal e fração de ejeção menor que 50% e/ou *strain* global longitudinal maior do que 15%. O período de avaliação dos parâmetros ecocardiográficos foi de 2 anos após o ecocardiograma realizado antes do tratamento.

#### 3.2 Material

Foi utilizado o aparelho de ecocardiografia da marca GE, modelo VIVID 9, com transdutor setorial eletrônico multifrequencial (2,5 a 3,5 MHz) com recurso de Doppler pulsado, Doppler contínuo, Doppler tecidual, *speckle tracking*, mapeamento de fluxo a cores e monitorização eletrocardiográfica contínua. As imagens foram armazenadas para calcular os parâmetros a posteriori no programa *EchoPac*, versão 202.

### **3.3 Casuística**

A amostra do presente estudo foi oriunda do universo de pacientes com câncer com seguimento mínimo de 24 meses na pesquisa piloto, perfazendo um total de 101 pacientes, que finalizaram tratamento com doxorubicina (DOX) entre fevereiro de 2010 a junho de 2011, com acompanhamento clínico e ecocardiográfico, sendo que 58 pacientes apresentam dados clínicos e ecocardiográficos completos.

Os ecocardiogramas foram realizados no serviço de ecocardiografia do Hospital Dom Pedro de Alcântara/Santa Casa de Misericórdia de Feira de Santana (HDKA), Bahia, por profissionais com títulos de especialista em cardiologia e ecocardiografia pela Sociedade Brasileira de Cardiologia. Os ecocardiogramas foram realizados seguindo-se as recomendações da Sociedade Americana de Ecocardiografia. As imagens foram armazenadas em HD externo para análise de um examinador local (UFPB) e avaliadas pelo uso do programa *EchoPac* versão 202, para cálculos dos parâmetros de interesse.

Todos os pacientes preencheram os critérios de inclusão no estudo piloto. Eles foram meticulosamente informados acerca da pesquisa e assinaram um Termo de Consentimento Livre e Esclarecido aplicado pelos pesquisadores.

### **3.4 Critérios de inclusão**

Os critérios de inclusão foram: idade maior ou igual a 18 anos, ter o diagnóstico inicial de câncer de mama e estar realizando acompanhamento com oncologista no HDKA-UNACON (Unidade de Alta Complexidade em Oncologia), ecocardiograma e eletrocardiograma com boa qualidade técnica para análise e ter seguido o protocolo do estudo piloto por pelo menos 24 meses.

### **3.5 Critérios de exclusão**

Os critérios de exclusão foram: recusa do paciente em participar da pesquisa, hipertensão arterial sistêmica não controlada, doença valvar cardíaca moderada ou grave, complexo QRS alargado no eletrocardiograma basal, outro ritmo que não o sinusal no eletrocardiograma basal, sorologia positiva para doença de Chagas,

FEVE < 52%, história prévia de insuficiência cardíaca ou doença arterial coronariana, ecocardiograma de má qualidade técnica, perda de seguimento precoce que não possa ter nenhum parâmetro avaliado, retirada do TCLE inicial e não ter seguimento mínimo de 24 meses, excluindo-se os casos de óbito.

### **3.6 Exame ecocardiográfico**

O ecocardiograma transtorácico foi realizado no aparelho Vivid 9 (General Electric-GE), por um único examinador ecocardiografista com certificado de habilitação em ecocardiografia pelo Departamento de Imagem Cardiovascular da Sociedade Brasileira de Cardiologia (DIC/SBC).

As imagens foram adquiridas em decúbito lateral esquerdo ou decúbito dorsal conforme as recomendações da Sociedade Americana de Ecocardiografia com cortes ecocardiográficos em posições subcostal, paraesternal, apical e supraesternal, utilizando o modo-M, bidimensional, além do Doppler pulsado, contínuo, em cores, além do Doppler tecidual. (21)

Foram avaliadas as seguintes imagens adquiridas com o modo bidimensional como as dimensões da raiz aorta em nível dos seios aórticos no final da diástole, do átrio esquerdo no final da sístole, todas no corte paraesternal eixo longo, diâmetros diastólico e sistólico do ventrículo esquerdo (DdVE e DsVE), espessuras diastólicas do septo interventricular e da parede posterior do ventrículo esquerdo, no corte paraesternal eixo longo. A FEVE foi estimada pelo método de Simpson biplano modificado.

O volume máximo do átrio esquerdo foi calculado no final da sístole nos cortes apicais 4 e 2 câmeras usando o método biplanar de discos, já o volume máximo do átrio direito foi calculado pelo mesmo método, no corte apical 4 câmeras. As medidas foram corrigidas pela área de superfície corporal.

Nas imagens adquiridas pelo modo-M foram calculadas a excursão sistólica do plano anular da valva tricúspide (TAPSE), utilizada para análise da função sistólica do ventrículo direito e a excursão sistólica do plano anular valvar mitral septal e lateral (MAPSE), utilizada para análise da função sistólica do ventrículo esquerdo. O Doppler

colorido foi utilizado para avaliar os fluxos, e a presença de estenose e insuficiência valvar.

A análise do fluxo transvalvar mitral foi feita utilizando o Doppler pulsado em corte apical 4 câmaras, colocando a amostra de volume na extremidade dos folhetos da valva mitral na diástole. A onda E corresponde à velocidade de enchimento diastólico precoce do ventrículo esquerdo, a onda A corresponde à velocidade de enchimento tardio do ventrículo esquerdo, durante a contração atrial, e dessa forma se obtém a relação E/A, que é um dos parâmetros utilizados para análise da função diastólica.

O Doppler tecidual no anel mitral septal, lateral, anterior e inferior, além da região lateral do anel tricúspideo foram utilizados para o cálculo das velocidades sistólica (S'), diastólica precoce (e') e tardia (a'). A relação E/e' é útil para estimativa não-invasiva da pressão de enchimento do ventrículo esquerdo e análise da função diastólica. (22)

Nos pacientes que apresentaram insuficiência tricúspide o Doppler contínuo foi empregado para calcular a velocidade máxima do refluxo tricúspideo. A medida da pressão sistólica do ventrículo direito é igual a pressão sistólica da artéria pulmonar. A fórmula pressão sistólica da artéria pulmonar=  $4 \times V^2$  + pressão do átrio direito, em que V é a velocidade máxima do refluxo tricúspide. A pressão do átrio direito foi estimada de acordo com o diâmetro da veia cava inferior.

Para a aquisição das imagens do SLG, o aparelho Vivid 9 foi ajustado para taxa de quadros entre 40 e 80 quadros por segundo, sendo conferido o sinal do eletrocardiograma. Cortes apicais de três câmeras, quatro câmeras e duas câmeras foram realizados para análise dos segmentos miocárdicos.

O índice de desempenho miocárdica foi obtido pela soma da contração isovolumétrica e o relaxamento isovolumétrico dividido pelo tempo total de ejeção.<sup>22</sup>

### **3.7 Critérios de cardiotoxicidade**

A definição de cardiotoxicidade utilizada neste estudo foi baseada pelo ecocardiograma com queda de dez pontos percentuais e uma FEVE menor do que

50%, associada a uma redução de SLG maior que 15% entre o primeiro e subsequentes momentos do estudo.

### **3.8 Aspectos éticos**

O planejamento do presente trabalho usou como referencial ético a Resolução 466/2012 da Comissão Nacional de Ética em Pesquisa do Ministério da Saúde. Os seguintes pontos merecem consideração especial por estarem diretamente ligados ao tipo de trabalho desenvolvido:

- O Projeto de Pesquisa Piloto foi aprovado pelo Comitê de Ética em Pesquisa (CEP) da Universidade Estadual de Feira de Santana - CAEE 02590212.0.0000.0053;

- As pessoas que preencherem os critérios de inclusão da pesquisa foram informadas meticulosamente acerca da mesma e, caso não concordassem em participar, deram o seu consentimento através da assinatura do Termo de Consentimento Livre e Esclarecido;

- Foi garantido que não haveria referência aos nomes nem ao número de matrícula dos pacientes nos respectivos serviços;

- Foi obtida a autorização do Diretor médico do HDPA-UNACON;

- Foi assegurado aos pacientes que, se os resultados da pesquisa mostrar alguma forma de benefício para eles, os mesmos seriam informados imediatamente;

- Os pacientes com diagnóstico de disfunção ventricular foram tratados conforme recomendações das diretrizes internacionais;

- Este estudo, como utiliza apenas os dados de imagens e desfechos do projeto piloto, recebeu carta de anuência do pesquisador Prof. Dr. André Luiz Cerqueira de Almeida para acesso aos dados coletados e armazenados, sendo analisados pelo pesquisador local Prof. Dr. Marcelo Dantas Tavares de Melo através do programa *EchoPac v202*;

- Este estudo foi submetido ao comitê de ética do Hospital Universitário Lauro Wanderley da Universidade Federal da Paraíba CAEE: 15775519.0.0000.5183 e ao comitê de ética da Faculdade de Medicina da Universidade de São Paulo FMUSP CAEE: 15775519.0.3001.0065.

## 4 ANÁLISE ESTATÍSTICA

Para o estudo foram fornecidas diferentes tabelas em formato de planilhas excel, cada uma com um conjunto específico de dados relacionados a um grupo de pacientes submetidas a tratamento de câncer de mama. Havia um total de duas tabelas com diferentes dados ecocardiográficos complementares. Ademais, havia uma tabela no formato de agenda em que se marcava uma cronologia da ocorrência de todos os eventos, bem como outra que explicava a causa de óbito de certos pacientes, sendo estas últimas não importantes para a base no total.

Assim, foram realizadas adaptações nas três tabelas de forma a tornar o nome dos pacientes como índice, ou seja, como coluna de referência. Tudo isso para se realizar a concatenação de todas as tabelas a partir de tal coluna, gerando como resultado final uma única tabela.

A avaliação de cada modelo proposto neste trabalho se deu em relação aos parâmetros de acurácia e à área sob a curva de característica de operação do receptor (ROC AUC). Ademais, a avaliação de todos os modelos aplicados individualmente também se dará de forma a se observar os indicadores estatísticos essenciais. (23)

Neste contexto, torna-se necessário conhecer alguns dos principais conceitos mais utilizados nesta etapa. Para isso, mostramos uma definição prática dos principais descritos:

- 1) Matriz de confusão: Trata-se de uma tabela que apresenta os erros e os acertos previstos por um modelo comparado com o que se verificou nos dados originais. Para tal designação existem os seguintes parâmetros: Verdadeiro positivo (VP), verdadeiro negativo (VN), falso positivo (FP), falso negativo (FN), como demonstrado na Tabela 3.

**Tabela 3** - Representação dos parâmetros na matriz de confusão

MATRIZ DE CONFUSÃO		Observados pelo modelo	
		Positivo	Negativo
Dados Reais	Positivo	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Negativo	Falso Positivo (FP)	Verdadeiro Negativo (VN)

Fonte: Produzido pelo autor.

- 2) Acurácia: trata-se de um medidor do desempenho geral do modelo. Para isso ele faz a relação da quantidade de classificações corretas pelo total de classificações existentes.

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN}$$

- 3) Precisão/Valor Preditivo Positivo: Avalia a relação entre a quantidade de classificações verdadeiras positivas existentes e as que o modelo classificou como positivas, podendo estar erradas ou não.

$$\text{Precisão} = \frac{VP}{VP + FP}$$

- 4) *Recall* (Sensibilidade): Avalia dentre todas as situações classificadas como positivas quantas estão de fato corretas se comparado aos dados reais.

$$\text{Recall} = \frac{VP}{VP + FN}$$

- 5) F1-Score: Média harmônica entre precisão e *recall*.

$$\text{F1-Score} = \frac{2 * \text{Precisão} * \text{Recall}}{\text{Precisão} + \text{Recall}}$$

Após a limpeza dos dados foi utilizada a biblioteca *PyCaret* para a rápida visualização do desempenho de vários modelos de aprendizado de máquina. Para o escopo do projeto foram escolhidos os modelos de floresta aleatória e árvore de decisão, visto que além de serem úteis para realizar previsões podem indicar grau de relevância de diversos atributos avaliados na base de dados, o que se torna útil para determinar variáveis mais significativas dentro do contexto de saúde.

Após a definição do modelo escolhido, a biblioteca do *SciKit Learn* foi escolhida para a execução do modelo, obtendo primeiro baixos valores de *recall* (sensibilidade) para o desfecho negativo, seguiu-se então com um *GridSearch* de parâmetros buscando a maximização dessa métrica, obteve-se então melhores resultados.

## 5 Resultados

De uma amostra inicial de 101 pacientes, no final do estudo obteve-se dados de 58 pacientes do sexo feminino com média de  $52,49 \pm 12,97$  anos, índice de massa corpórea de  $25,27 \pm 8,3$  por metro<sup>2</sup>, como apresentado na Tabela 2. Dentre os atributos existentes nas tabelas, compreendeu-se dados socioeconômicos, uso de diversas medicações, sintomáticos até de medicações específicas da terapia do câncer, além de parâmetros ecocardiográficos durante as cinco fases do estudo.

**Tabela 2** - Características clínicas e demográficas de todas as pacientes antes do início do tratamento quimioterápico

Variável	Amostra Total n=58
FC (bpm)	78,5 [63; 120]
HAS	25 (42,6%)
DM	4 (6,6%)
DAC	0 (0%)
Tabagismo	3 (4,9%)
DLP	5 (8,2%)
Etilismo	9 (14,8%)
Obesidade	4 (6,6%)
Sedentarismo	43 (73,8%)
Uso de IECA	8 (13%)
Uso de BB	1 (1,6%)
Uso de BRA	8 (13%)

Radioterapia	45(77%)
Transtuzumab	13 (22,9%)

Legenda: ASC: área de superfície corporal, BB: betabloqueadores, BRA: bloqueadores do receptor da angiotensina, Circ.Abd: circunferência abdominal, DAC: doença arterial coronariana, DLP: dislipidemia, DM: diabetes mellitus tipo 2, FC: frequência cardíaca, HAS: hipertensão arterial sistêmica, IECA: inibidor da enzima conversora de angiotensina, IMC: índice de massa corporal, PAD: pressão arterial diastólica, PAS: pressão arterial sistólica.

Nota: As variáveis numéricas estão apresentadas como mediana e amplitude. A área de superfície corporal foi calculada pela fórmula de Dubois ( $BSA = 0.007184 * Altura^{0.725} * Peso^{0.425}$ ).

No atual estudo, nove pacientes apresentaram queda de dez pontos percentuais e uma FEVE menor do que 50% no último momento, somaram-se oito com redução do SLG maior que 15% entre o primeiro e subsequentes momentos do estudo, totalizando 17 pacientes agora definidos como cardiotoxicidade. Associado aos desfechos de óbito (12 pacientes) e/ou cardiotoxicidade no momento final (17 pacientes), totalizando 29 pacientes.

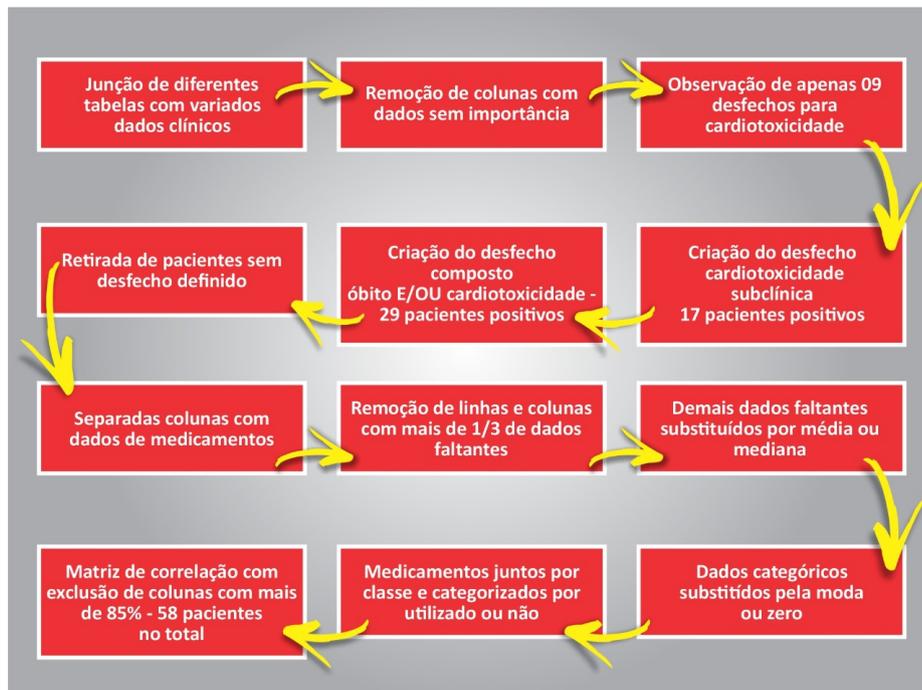
O primeiro método de limpeza dos dados faltantes foi a retirada de pacientes (linhas) que não possuíam desfecho definido, seguindo foram separadas as colunas

Para maior eficiência do modelo, as linhas e colunas com mais de 90% dos dados faltantes foram removidos, valor escolhido para maximizar a quantidade de pacientes e buscar um valor razoável e pequeno de imputações. Para os dados faltantes restantes foram utilizadas duas metodologias: a primeira foi a imputação de dados numéricos por média e a segunda por mediana, considerando a curva distribuição destes e a proximidade ou não a uma curva normal, respectivamente. Para os dados categóricos foram imputados por moda ou pelo valor zero. Além disso, foram excluídas colunas que não possuíam dados relacionados com a pesquisa, como a coluna de ID que trazia número de identificação de cada paciente.

Um último passo foi a visualização da matriz de correlação dos dados, que nos mostra o quanto determinadas variáveis apresentam um comportamento correlacionado e definido por uma regra matemática com boa aproximação, o que não se torna algo interessante, visto que se deve buscar o máximo de independência entre os dados. Com isso, foi possível a retirada daqueles com valor de correlação a partir

de 85%. Bem como a retirada dos atributos utilizados para a definição de alvos, como a SLG. Totalizou-se por fim 58 pacientes e 235 atributos, como demonstrado na Figura

**Figura 5** - Fluxograma com as etapas de pré-processamento dos dados.



**Fonte:** Elaborado pelo autor

Neste contexto foram totalizados os seguintes grupos de pacientes:

- Presença de cardiotoxicidade clínica e/ou óbito ao final das cinco avaliações: 29 pacientes
- Ausência de qualquer um dos desfechos: 29 pacientes

Após a limpeza dos dados foi utilizada a biblioteca *PyCaret* para a rápida visualização de desempenho de vários modelos de aprendizado de máquina. A floresta aleatória e árvore de decisão foram selecionados, visto que além de serem úteis para realizar previsões, podem indicar grau de relevância de diversos atributos avaliados na base de dados, o que se torna útil para determinar variáveis mais significativas dentro do contexto de saúde. Após a definição do modelo escolhido, a biblioteca do *SciKit Learn* foi selecionada para a execução do modelo, obtendo

primeiro baixos valores de *recall* (sensibilidade) para o desfecho negativo, seguiu-se então com um *GridSearch* de parâmetros buscando a maximização dessa métrica, obteve-se então melhores resultados.

Portanto, a Tabela 4 apresenta o resultado da etapa inicial para escolha dos melhores modelos, colocando em ordem de importância de acordo com os parâmetros mencionados anteriormente.

**Tabela 4-** Teste global com principais parâmetros na base de dados a partir de diferentes modelos de aprendizado de máquina

<b>Modelo</b>	<b>Accuracy</b>	<b>AUC</b>	<b>Recall</b>	<b>Precision</b>	<b>F1</b>
<b>K Neighbors Classifier</b>	<b>0.7509</b>	<b>0.7472</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
<b>Random Forest Classifier</b>	<b>0.7509</b>	<b>0.6736</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
<b>Cat Boost Classifier</b>	<b>0.7509</b>	<b>0.6556</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
<b>SVM - Radial Kernel</b>	<b>0.7509</b>	<b>0.6111</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
<b>AdaBoost Classifier</b>	<b>0.7509</b>	<b>0.6278</b>	<b>0.1111</b>	<b>0.1667</b>	<b>0.1333</b>
<b>Light Gradient Boosting Machine</b>	<b>0.7509</b>	<b>0.5000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
<b>Dummy Classifier</b>	<b>0.7509</b>	<b>0.5000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>

**Fonte:** Produzido pelo autor

Logo, observa-se que os algoritmos que tiveram os maiores valores de acurácia e área sob a curva de característica de operação do receptor foram o *K Neighbors classifier* e o floresta aleatória *classifier*, sendo tais modelos os escolhidos para a realização das avaliações descritas neste trabalho.

Por fim, para iniciar o treinamento e os testes da base de dados com os respectivos algoritmos de aprendizado de máquina selecionados, 30% do conjunto de dados foi separado para teste (18 pacientes), portanto os 40 pacientes restantes foram divididos em 3 camadas, que tem como objetivo para o treinamento do modelo utilizando a estratégia *StratifiedKFold*, e logo após foram executados os modelos.

Em seguida uma segunda metodologia foi aplicada neste trabalho, que envolveu aprimoramento do conjunto de dados com nova limpeza de dados e formulação de variáveis diferentes das já então estabelecidas, bem como uso direto do algoritmo *XGB classifier*.

Após a etapa de pré-processamento e limpeza dos dados, realizou-se a comparação entre modelos de classificação consagrados, utilizando as bibliotecas *Scikit-Learn* e *XGBoost*. Para tanto, foram implementados modelos de aprendizado de máquina com técnica de *ensemble* visando à otimização dos resultados, empregando-se os algoritmos floresta aleatória *classifier* e *XGBoost classifier*. A validação cruzada *StratifiedKFold* foi utilizada com valores de  $k$  iguais a 3 e 5, considerando a limitação de dados do conjunto de dados em questão.

. Em cada modelo selecionado, as bibliotecas *Scikit-Learn* e *Scikit-Optimize* foram empregadas para a otimização dos *hiper parâmetros*, com foco na maximização da área sob a curva de característica de operação do receptor numa amostra separada de teste compondo 20% do conjunto de dados.

### **5.1 Técnica de avaliação de melhores parâmetros dos modelos**

Após a otimização dos hiper parâmetros, os modelos foram treinados, e a partir deles, foram gerados gráficos de importância de *features* e valores *Shap*, com o objetivo de entender a relevância de cada variável na predição do modelo.

A importância de variável (*feature importance*) é uma métrica que busca quantificar a contribuição individual de cada variável na previsão do modelo. Estes valores são obtidos por técnicas de permutação, com base na análise da frequência com que cada variável é utilizada para dividir os dados, nos exemplos que usam árvores de decisão, e em modelos lineares os coeficientes de regressão representam a importância de cada variável na previsão.

Os valores Shap (*SHapley Additive exPlanations*), através da biblioteca homônima são uma técnica mais recente e avançada para explicar as previsões de modelos de aprendizado de máquina. Ao contrário do *feature importance*, que fornece uma medida global da importância de cada variável, os valores Shap fornecem uma explicação local, ou seja, indicam como cada variável contribuiu para a previsão em um caso específico, utilizando um método baseado em teoria dos jogos para calcular a contribuição de cada variável na previsão, levando em consideração as interações entre as variáveis.

Para tal interpretação foram utilizados dois tipos de gráficos: O *summary plot* que corresponde ao gráfico que avalia o quanto cada variável gera de importância, tanto para mais, quanto para menos, em relação ao desfecho, a partir do valor que é entregue. E o gráfico de média absoluta do *SHAP*, que indica o quanto cada variável influenciou globalmente nos valores finais dos desfechos

## **5.2 Resultados para escolha do classificador**

A seguir serão descritos os resultados obtidos após a execução dos melhores modelos, sendo eles o *K Neighbors classifier* e o floresta aleatória *classifier*. Neste capítulo também será apresentado uma análise das variáveis mais importantes a partir de cada modelo.

### **5.2.1 Predição por Classificadores**

Após a primeira seleção, cada modelo foi executado individualmente e otimizado para melhor classificar os dados limpos. Nesta seção serão expostas as pontuações de teste, com dados separados da instância de treinamento, assegurando

a generalização do modelo escolhido. Para cada aplicação, foi avaliado o modelo antes e depois da otimização dos resultados a partir da biblioteca *Grid Search*. Dessa forma, foi possível analisar comparativamente cada situação.

### 5.2.1.1 Classificador *K-Nearest Neighbors*

. Antes da Otimização, o modelo possuía uma média maior de acurácia de 75,09% e área sob a curva de característica de operação do receptor de 74,72%, porém não conseguia discriminar o defeito, pois o *recall*, a precisão e, obviamente a F1 eram iguais a 0, como pode ser visto na Tabela 5 abaixo.

**Tabela 5** - Métricas de avaliação inicial do modelo *K Neighbors* antes do processo de otimização

	<b>Acurácia</b>	<b>AUC</b>	<b>Recall</b>	<b>Precisão</b>	<b>F1</b>
Fold 0	0.7143	0.8250	0.0000	0.0000	0.0000
Fold 1	0.7692	0.9000	0.0000	0.0000	0.0000
Fold 2	0.7692	0.5167	0.0000	0.0000	0.0000
<b>Média</b>	<b>0.7509</b>	<b>0.7472</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
DP	0.0259	0.1659	0.0000	0.0000	0.0000

**Fonte:** Produzida pelo autor

.Portanto, foi executado a otimização de forma que o modelo foi rodado em duas etapas, uma considerando apenas os *folds* e por fim aplicando-se a base de teste separada anteriormente, visto que o objetivo é verificar como a base de teste teria sua classificação apresentada após o treinamento do modelo. Com isso, foi possível obter 25% de *recall* e 100% de precisão:

**Tabela 6** - Métricas de avaliação do modelo *K Neighbors* após o processo de otimização.

	<b>Acurácia</b>	<b>AUC</b>	<b>Recall</b>	<b>Precisão</b>	<b>F1</b>
Fold 0	0.7857	0.6250	0.2500	1.0000	0.4000
Fold 1	0.6923	0.4500	0.0000	0.0000	0.0000
Fold 2	0.6923	0.4500	0.0000	0.0000	0.0000
<b>Média</b>	<b>0.7234</b>	<b>0.5083</b>	<b>0.0833</b>	<b>0.3333</b>	<b>0.1333</b>
DP	0.0440	0.0825	0.1179	0.4714	0.1886

Fonte: Produzido pelo autor

Por fim, o modelo alcançou 40% de *recall* e 50% de precisão ao se utilizar o conjunto de dados do teste. Na Tabela 7 abaixo estão apresentados os dados na nova implementação logo após a otimização com cada métrica avaliada na base de dados de teste.

**Tabela 7** - Métricas de avaliação do modelo *K Neighbors* no conjunto de dados de teste

<b>Modelo</b>	<b>Acurácia</b>	<b>AUC</b>	<b>Recall</b>	<b>Precisão</b>	<b>F1</b>
KNN	0.7222	0.6231	0.4000	0.5000	0.4444

Fonte: Produzida pelo autor

O processo de otimização envolve a escolha de diversos parâmetros disponíveis dentro da biblioteca de cada modelo de aprendizado de máquina, chamados estes de *Hyper Parameters*. A escolha de tais parâmetros pode variar de acordo com os dados, disposição, quantidade, organização e objetivo da base. Desta forma, a biblioteca *GridSearch* após realizar uma sequência de testes, apontou os seguintes parâmetros como mais relevantes para o modelo descrito na Tabela 7

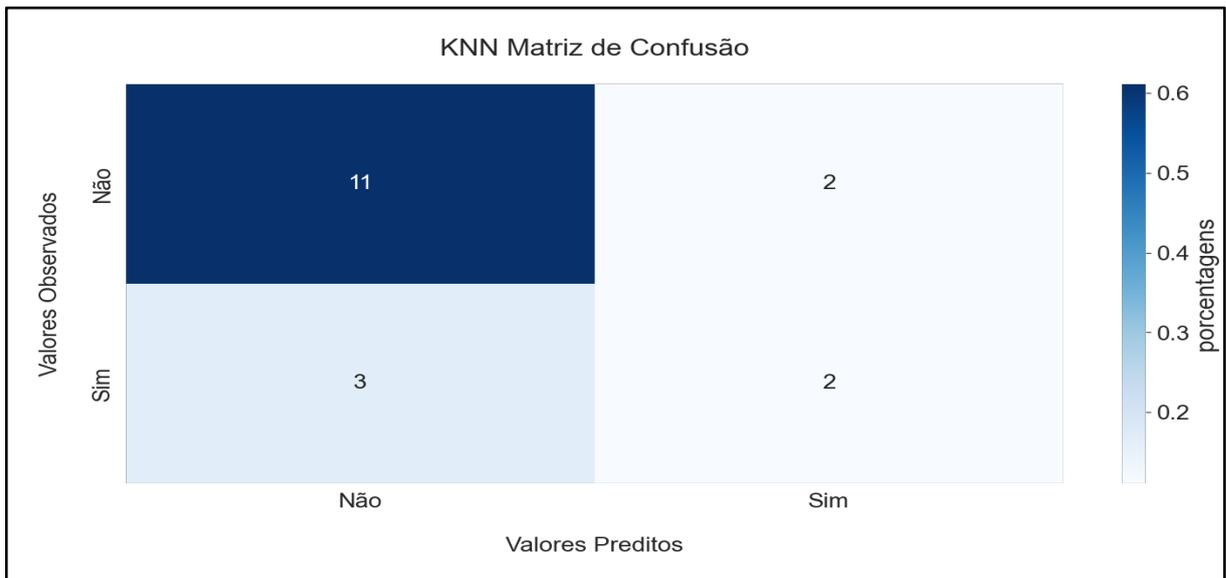
**Tabela 8** - Hiper Parâmetros mais relevantes encontrados pela otimização da biblioteca *GridSearchCV* para o *K Neighbors Classifier*.

Nome do modelo	KNeighborsClassifier
Métrica otimizada	Recall
Algoritmo	auto
leaf_size	30
metric	minkowski
metric_params	None
n_neighbors	1
p	2
weights	uniform

**Fonte:** Produzido pelo autor

Observa-se, portanto, que o modelo a partir do conjunto de dados do teste possui 72,34% de acurácia após a otimização, entretanto, apresenta 40% de *recall* e 50% de precisão. Ao se analisar o peso que dados falsos negativos possuem para uma predição em avaliação na área da saúde, esperava-se ter um *recall* com valores mais elevados nestes modelos, visto que se deve diminuir ao máximo os falsos negativos. Entretanto, tais resultados ainda podem ser considerados pouco relevantes, visto que o tamanho da base de dados é fundamental para se obter melhores observações, todavia, como a base utilizada ainda é considerada pequena, houve poucos dados para melhorar a capacidade do modelo de identificar os verdadeiros positivos e diminuir os falsos positivos.

**Figura 6** - Matriz de confusão obtida pelo *K Neighbors classifier*

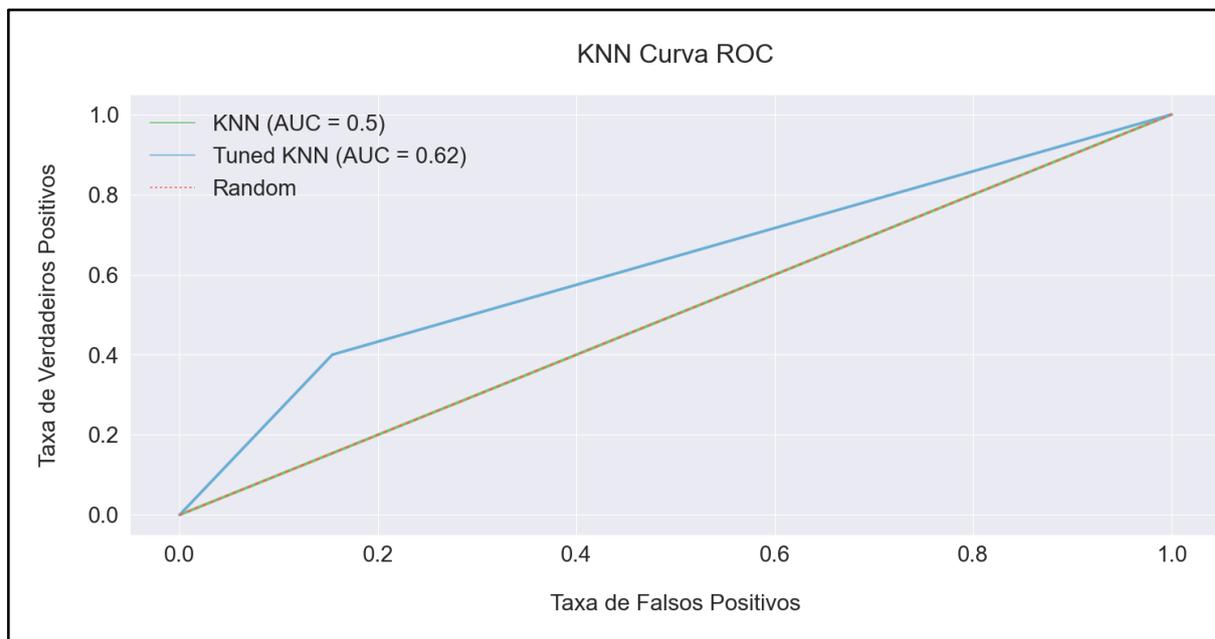


**Fonte:** Produzido pelo autor

Ao se analisar a matriz de confusão da Figura 6, observa-se com mais exatidão os valores alcançados pelas métricas, além da pouca quantidade de dados disponíveis para se realizar o teste. Entretanto, pode-se avaliar pela matriz que mesmo com uma pequena quantidade de dados, o modelo direcionou a apresentar possibilidades de previsões.

Na Figura 7 abaixo observa-se a área sob a curva de característica de operação do receptor para o modelo avaliado de 62%.

**Figura 7** - Área sob a curva de característica de operação do receptor para o modelo KNN



**Fonte:** Produzido pelo autor

### 5.2.1.2 Classificador floresta aleatória

O classificador floresta aleatória foi apresentado como o segundo modelo com maior acurácia 75,09% e área sob a curva de característica de operação do receptor de 65,36%. Neste contexto, foi realizado o mesmo esquema de testes observado anteriormente para o KNN. Abaixo, podemos observar as métricas de avaliação do modelo antes da otimização pelo *GridSearchCV* e após

**Tabela 9** - Métricas de avaliação do modelo Random Forest antes do processo de otimização.

	<b>Acurácia</b>	<b>AUC</b>	<b>Recall</b>	<b>Precisão</b>	<b>F1</b>
<b>Fold 0</b>	0.7143	0.7375	0.0000	0.0000	0.0000
<b>Fold 1</b>	0.7692	0.7333	0.0000	0.0000	0.0000
<b>Fold 2</b>	0.7692	0.5500	0.0000	0.0000	0.0000
<b>Média</b>	<b>0.7509</b>	<b>0.6736</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
DP	0.0259	0.0874	0.0000	0.0000	0.0000

Fonte: Produzida pelo autor

Mais uma vez, foi executada a otimização do modelo, sendo rodado em duas etapas, uma considerando apenas os *folds* no treinamento e, por fim, aplicando-se a base de teste separada anteriormente. Com isso, foi possível obter mais uma vez 25% de *recall* e 100% de precisão em um dos *folds*, semelhante ao que havia sido observado pelo KNN, o que mostra que os algoritmos alcançaram resultados muito semelhantes em ambas as aplicações.

**Tabela 10** - Métricas de avaliação do modelo *floresta* aleatória após o processo de otimização

	<b>Acurácia</b>	<b>AUC</b>	<b>Recall</b>	<b>Precisão</b>	<b>F1</b>
Fold 0	0.7857	0.8000	0.2500	1.0000	0.4000
Fold 1	0.7692	0.9000	0.3333	0.5000	0.4000
Fold 2	0.6923	0.6000	0.0000	0.0000	0.0000
<b>Média</b>	<b>0.7491</b>	<b>0.7667</b>	<b>0.1944</b>	<b>0.5000</b>	<b>0.2667</b>
DP	0.0407	0.1247	0.1416	0.4082	0.1886

Fonte: Produzido pelo autor.

Por fim, o modelo alcançou 80% de *recall* e 57% de precisão ao se utilizar o conjunto de dados de teste. Na tabela xyz estão apresentados os dados na nova implementação, logo após a otimização pelo *GridSearchCV*, como apresentado na Tabela 11.

**Tabela 11** - Métricas de avaliação do modelo *floresta* aleatória no conjunto de dados do teste

<b>Modelo</b>	<b>Acurácia</b>	<b>AUC</b>	<b>Recall</b>	<b>Precisão</b>	<b>F1</b>	<b>Kappa</b>	<b>MCC</b>
RF	0.7778	0.8769	0.8000	0.5714	0.6667	0.5068	0.5230

Fonte: Produzida pelo autor

Na Tabela 12 observa-se os hiper parâmetros selecionados pelo *GridSearchCV* para o modelo floresta aleatória.

**Tabela 12** - Hiper Parâmetros mais relevantes encontrados pela otimização da biblioteca *GridSearchCV* para a floresta aleatória *classifier*

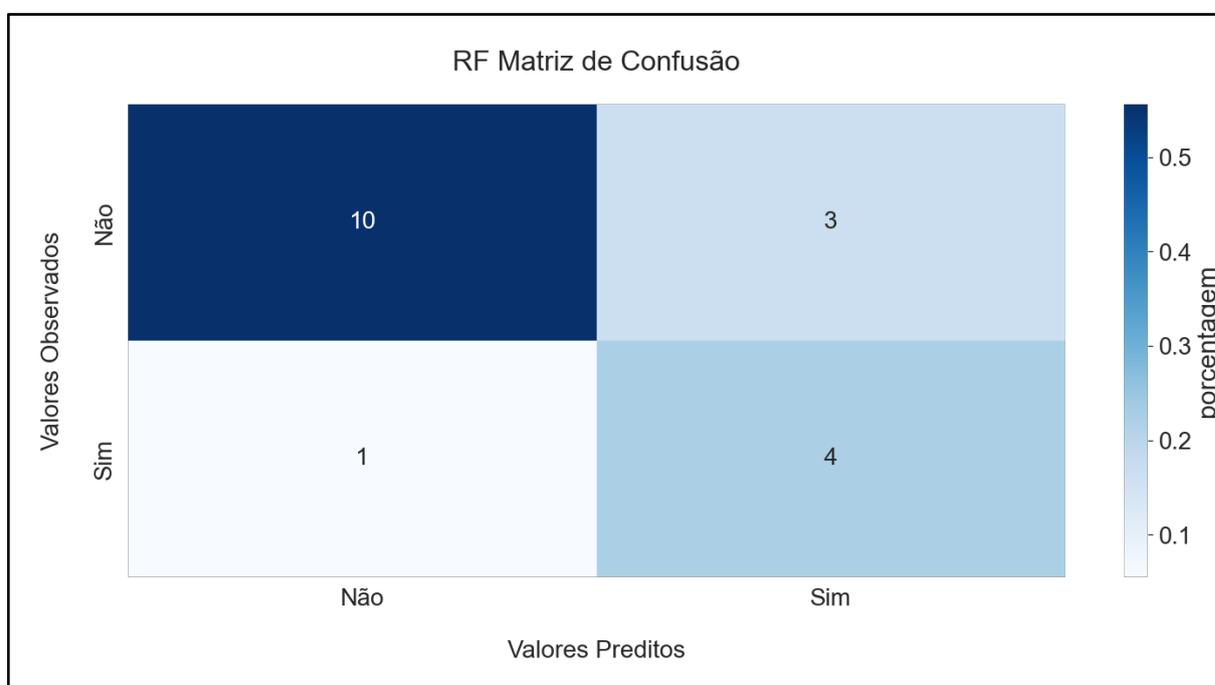
Nome do Modelo	<i>Floresta aleatória Classifier</i>
Optimize Metric	AUC
bootstrap	True
ccp_alpha	0.0
class_weight	balanced
criterion	gini
max_depth	8
max_features	1.0
max_leaf_nodes	None
max_samples	None
min_impurity_decrease	0.0001
min_samples_leaf	6
min_samples_split	7
min_weight_fraction_leaf	0.0
n_estimators	20
oob_score	False
Random_State	123
warm_start	False

Fonte: Produzida pelo autor

Neste contexto, observa-se que o modelo floresta aleatória *classifier* alcançou o melhor resultado do estudo a partir do conjunto de dados do teste. Os valores alcançados foram de 77,78% de acurácia e 87,89% de área sob a curva de característica de operação do receptor após a otimização. Portanto, pelo valor encontrado, percebe-se que o modelo foi muito eficaz ao determinar os valores verdadeiros positivos em detrimento dos falsos positivos, o que pode ser ainda melhor

constatado ao se observar os resultados alcançados pelo *recall*, que foi de 80%. O modelo ainda teve uma precisão de 57,14%, o que mostra uma dificuldade relacionada com a pequena quantidade de dados, entretanto, para esta aplicação, bem como para diversas aplicações em saúde, o valor do *recall* ainda continua sendo o mais desejado, visto que os riscos apresentados por não se detectar um paciente com alterações patológicas em doenças potencialmente letais é muito alto, o que pode ser evitado com valores de *recall* elevados, como visto na Figura 8.

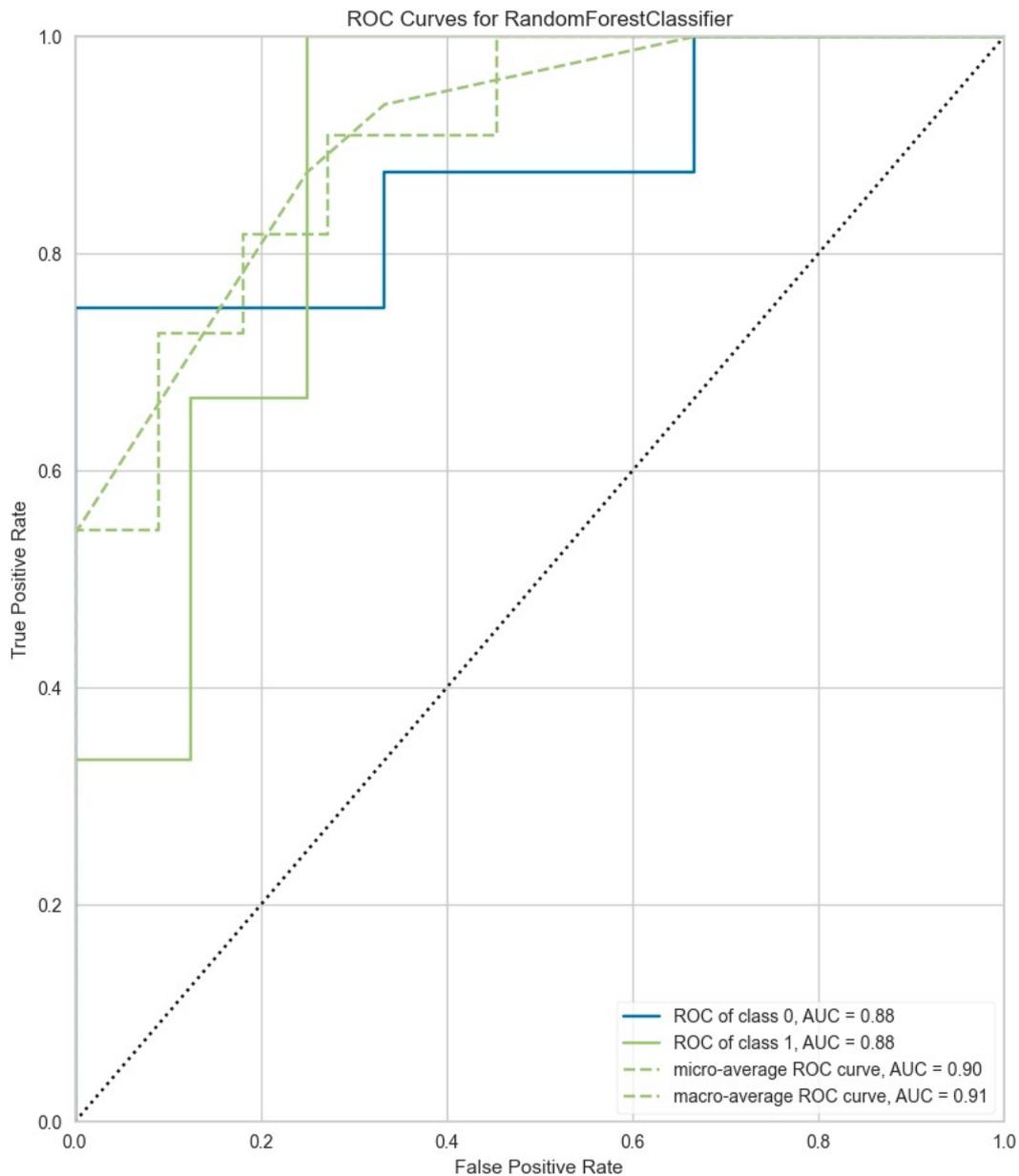
**Figura 8** - Matriz de confusão para o modelo floresta aleatória *classifier*



**Fonte:** Produzido pelo autor

Logo abaixo, na Figura 9, observa-se a área sob a curva de característica de operação do receptor em que se tem que por volta de 80% de verdadeiros positivos é apresentada uma taxa de pouco mais de 20% de falsos positivos e, após isso, a curva tende a um crescimento mais acelerado para falsos positivos, o que mostra que o modelo tem um comportamento considerado bom.

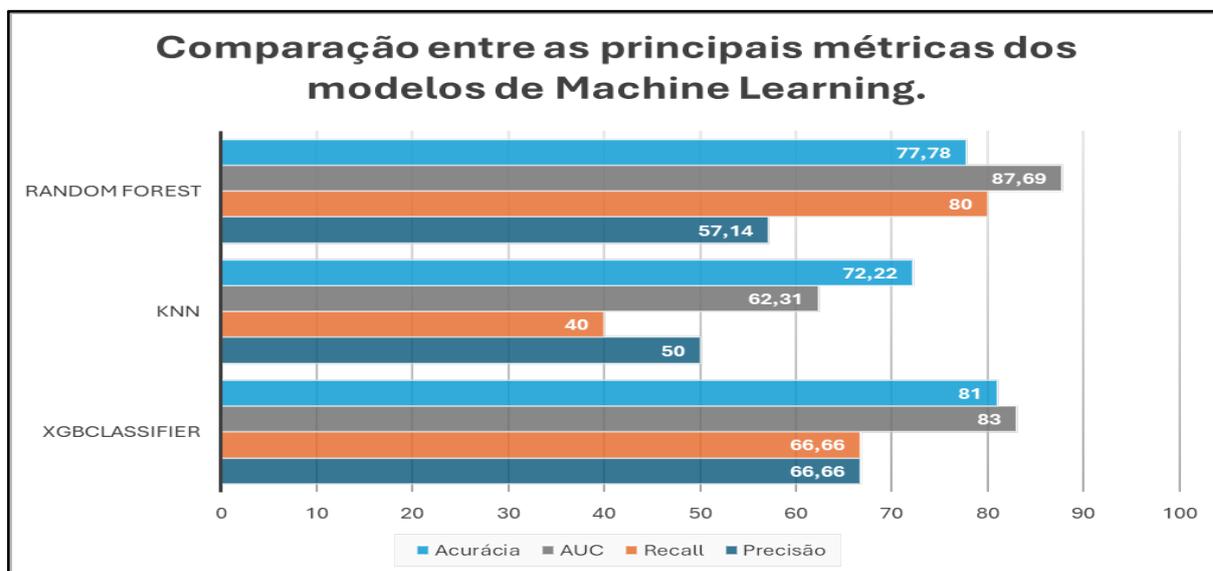
**Figura 9** - Área sob a curva de característica de operação do receptor para o modelo floresta aleatória *classifier*



**Fonte:** Produzida pelo autor

Abaixo é observado um gráfico com os principais valores alcançados de acurácia, área sob a curva de característica de operação do receptor, *recall* e precisão.

**Figura 10-** Comparação da Acurácia, AUC ROC, Recall e Precisão dos modelos de predição, em porcentagem



**Fonte:** Produzido pelo autor

#### 4.1.2 Análise das Variáveis

Neste subcapítulo serão apresentadas as variáveis/*features* que foram consideradas mais relevantes para o modelo selecionado floresta aleatória, considerando sempre o mesmo desfecho apresentado no decorrer deste trabalho.

##### 4.1.2.1 Análise por floresta aleatória

. No gráfico gerado pela análise SHAP do modelo final (Figura 10), temos em destaque as variáveis FEVE, onda e lateral (e lateral) TAPSE, índice de desempenho miocárdico e MAPSE, sendo a primeira a mais importante para discriminar os resultados neste modelo. Para compreender melhor as principais variáveis descritas para o modelo floresta aleatória pode-se observar o significado de cada uma delas na próxima Tabela 13.

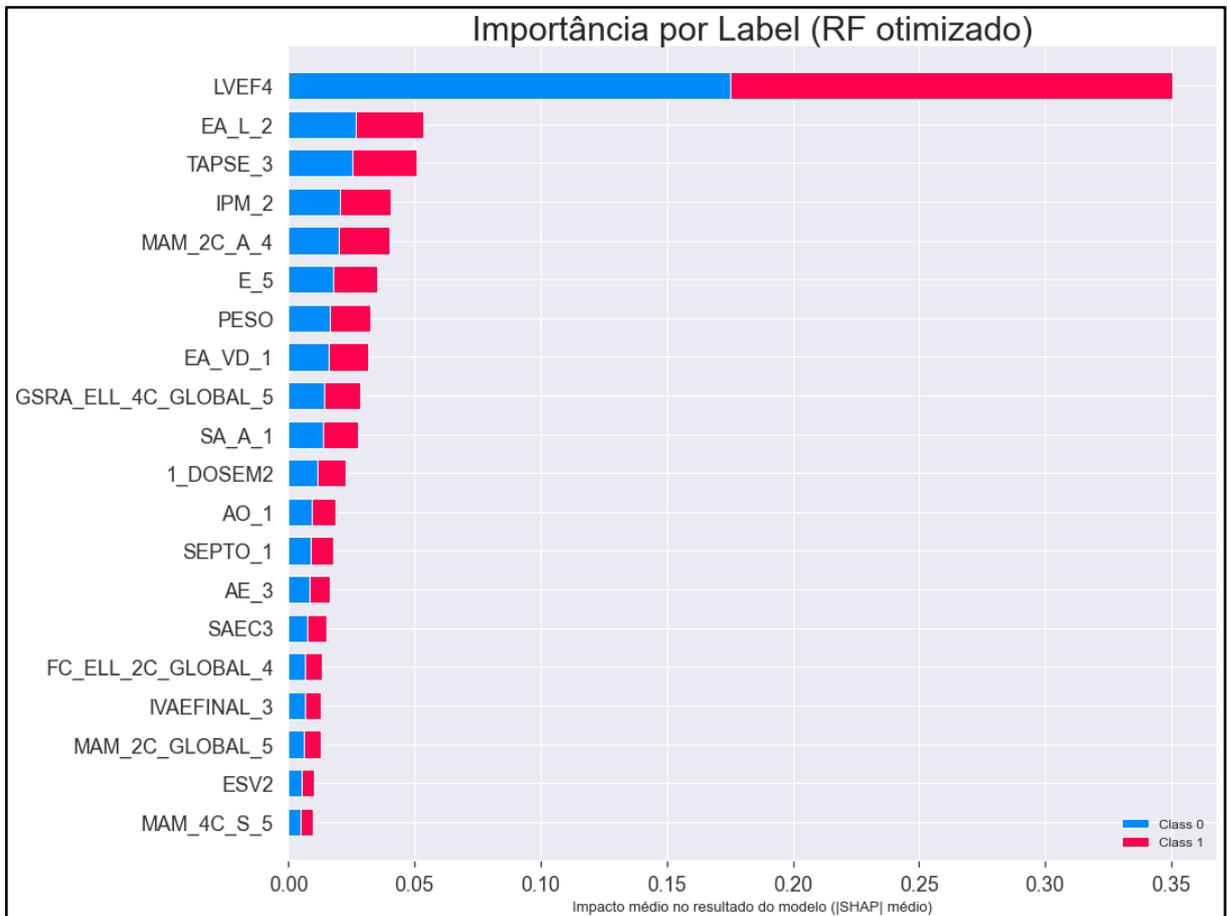
**Tabela 13** - Significado das principais variáveis encontradas a partir do modelo floresta aleatória

Variáveis observadas pela floresta aleatória	
Nome	Descrição
<b>FEVE</b> (LVEF4)	Fração de ejeção do ventrículo esquerdo
<b>e'</b> (EA_L_2)	Onda e' lateral (velocidade diastólica)
<b>TAPSE</b> (TAPSE_3)	Excursão sistólica de plano anular tricúspide
<b>IPM</b> (IPM_2)	Índice de desempenho miocárdico
<b>MAPSE</b> (MAM_2C_A_4)	Excursão sistólica do plano anular mitral
<b>E</b> (E_5)	Onda E do influxo mitral
<b>PESO</b>	Peso medido por paciente do estudo.
<b>e'VD</b> (EA_VD_1)	Onda e' da região lateral da valva tricúspide
<b>S'</b> (SA_A_1)	Velocidade sistólica

**Fonte:** Produzida pelo autor.

. Na Figura 11, pode-se observar o valor médio absoluto de importância para o modelo floresta aleatória, que cada uma das principais variáveis alcançou em relação às demais.

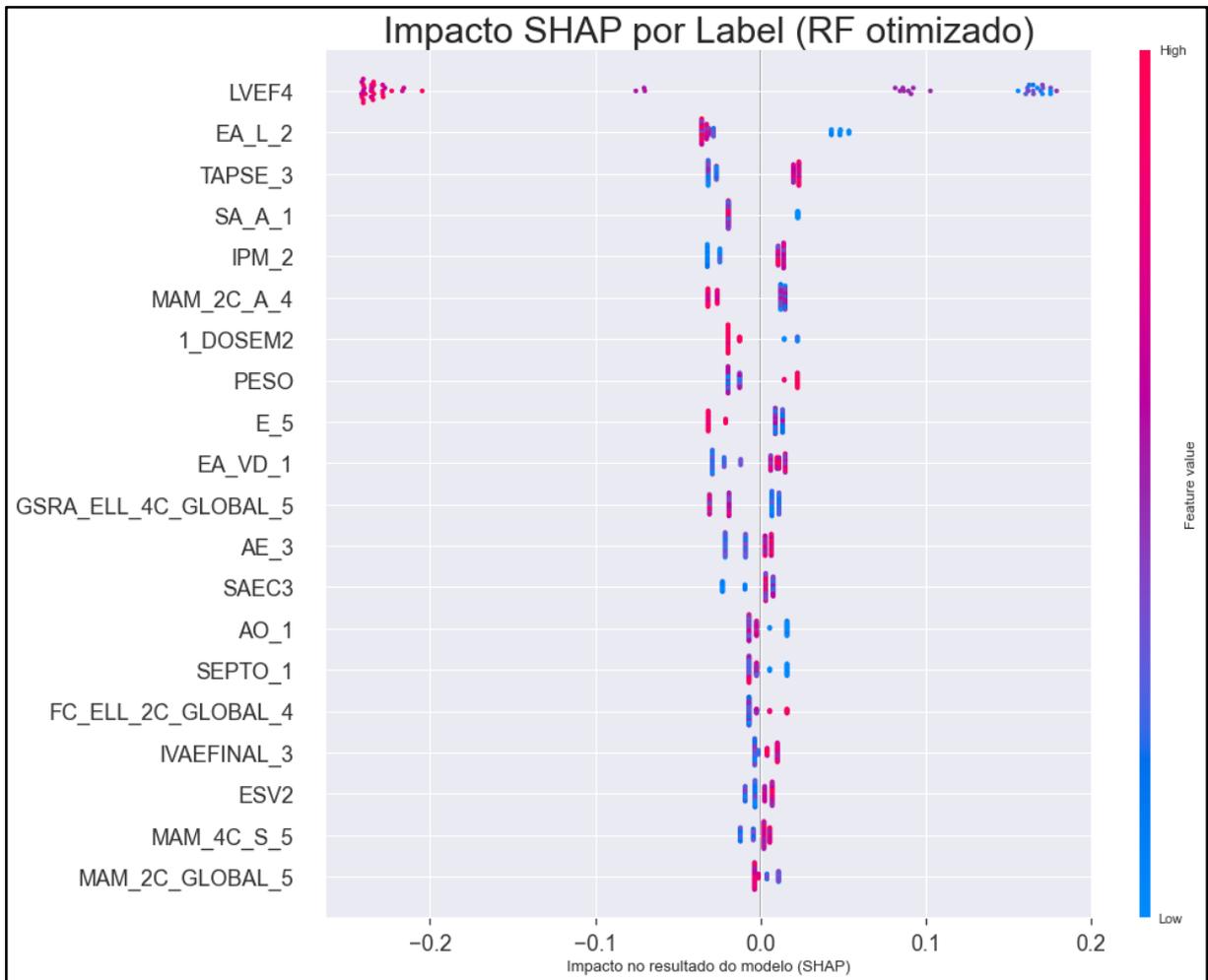
**Figura 11-** Gráfico da média absoluta de importância de cada variável pelo modelo SHAP para a floresta aleatória



**Fonte:** Produzida pelo autor.

Na Figura 12 temos um gráfico de pontos, onde cada ponto representa uma observação, a cor representa a distância do valor observado a média/mediana da variável, e a distância do eixo indica a importância que ela teve na construção do resultado do modelo floresta aleatória.

**Figura 12** - *Summary plot* no modelo SHAP para análise de importância de variáveis a partir da floresta aleatória



**Fonte:** Produzido pelo autor.

Ao se analisar o *plot* da Figura 11 também pode-se notar no gráfico de pontos, onde para FEVE, nesse modelo, valores mais altos (mais vermelhos) impactaram negativamente o modelo, ou seja, reduziram a chance de se haver desfecho patológico cardíaco. enquanto valores mais baixos tiveram um alto impacto positivo na discriminação do desfecho, ou seja, quanto menor o valor da FEVE, mais fácil foi ao modelo em classificar quem os pacientes com efeito patológico cardíaco.

## 5 DISCUSSÃO

A inteligência artificial já é uma realidade em várias áreas do conhecimento, com destaque especial na área cardiovascular, com avanços na eletrocardiografia, imagem cardiovascular, algorítmicos diagnósticos, que podem mudar a história natural de muitas patologias como as consequentes dos efeitos tóxicos do tratamento do câncer. (24)

### 5.1 Comparação entre os modelos preditivos

Durante o trabalho, foram criados dois modelos preditivos na metodologia (utilizando *K-Nearest Neighbors* e Floresta Aleatória) decididos a partir do método *Compare\_models* presente na biblioteca *Pycaret*, sendo esta biblioteca responsável por executar boa parte dos modelos com menor quantidade de linhas de código. Tal método informou uma lista com valores iniciais das principais métricas utilizadas para se avaliar globalmente alguns modelos, tendo as maiores notas para o KNN e a Floresta Aleatória.

O classificador floresta aleatória apresentou o melhor resultado do estudo com 77,78% de acurácia, 87,89% de área sob a curva ROC e 80% de *recall*, o método KNN mostrou 72,22% de acurácia, 62,31% de área sob a curva ROC e 40% de *recall*, enquanto o *XGBClassifier* apresentou 81% de acurácia, 83% de área sob a curva ROC e 66,6% de *recall*. Após a otimização as variáveis ecocardiográficas demonstradas pelos valores Shap (*SHapley Additive exPlanations*), por meio da biblioteca para explicar as previsões de modelos de aprendizado de máquina, foi observado que, além da redução da FEVE, índices como o TAPSE, velocidade da onda e' da região lateral da valva mitral, o índice de desempenho miocárdico e o MAPSE, apresentaram relação com cardiotoxicidade e/ou óbito.

Após ser realizada a escolha de tais algoritmos, cada um deles foi executado separadamente com a aplicação de uma biblioteca de otimização de resultados. Percebeu-se que presente estudo, o método Floresta Aleatória obteve os resultados de acurácia de 77,78% e 87,89% de área sob a curva de característica de operação do receptor, enquanto o método KNN 72,22% de acurácia e 62,31% de área sob a

curva de característica de operação do receptor. Por fim, o *XGBClassifier* apresentou acurácia de 81% de área sob a curva de característica de operação do receptor de 83%. Logo abaixo é observado um gráfico com os principais valores alcançados de acurácia, área sob a curva de característica de operação do receptor, *recall* e precisão.

Ao se comparar apenas os resultados de acurácia dos três modelos, observa-se que não existe uma grande diferença, entretanto, ao se observar a área sob a curva de característica de operação do receptor, nota-se que a floresta aleatória obteve um resultado mais expressivo na detecção de verdadeiros positivos e minimização de falsos positivos.

Como demonstrado, o classificador floresta aleatória apresentou o melhor valor quanto ao *recall* que foi de 80%, comparado a 62,31% de *recall* do KNN e 66,66% do *XGBClassifier*. Portanto, a capacidade da floresta aleatória em reduzir o valor de falsos negativos o tornou como o modelo mais adequado para a aplicação em questão, tudo isso mesmo com uma base de dados com número reduzido de pacientes.

Neste contexto, ao se observar tais métricas na Figura X compreende-se que o classificador floresta aleatória apresentou os melhores resultados globalmente o que indica um potencial para um algoritmo de boa resposta, mesmo em bases de dados pequenas e desbalanceadas.

## 5.2 Discussões acerca da análise de variáveis

Observa-se que diferentes variáveis se tornam importantes na capacidade de predição dos resultados em diferentes algoritmos com boas medidas de acurácia. Isto pode ser enxergado como um problema quando se tenta descobrir quais das variáveis teriam maior importância. Ao se observar a capacidade dos modelos apresentados em prever pacientes com efeitos cardíacos patológicos de cardiotoxicidade, combinado com o evento mais agressivo que seria o óbito, todos os modelos conseguiram prever tais desfechos. Entretanto, o modelo floresta aleatória apresentou potencial significativo nas previsões mesmo com uma quantidade reduzida de dados.

Neste contexto, ao se tentar relacionar o efeito que cada variável apresentou

em termos de importância, foi utilizado modelo de SHAP, que atribui uma importância que cada modelo tem, considerando todos os outros em diversas configurações possíveis e informa o quanto que o aumento ou a diminuição de determinado valor da variável pode influenciar no desfecho pretendido. Logo, todos os modelos de aprendizado de máquina apresentaram uma classificação de variáveis diferentes no grau de importância, entretanto com certa semelhança em alguns parâmetros.

O modelo floresta aleatória apresentou como variável principal FEVE, o que ao relacionar com o evento composto de cardiotoxicidade e/ou óbito, valores baixos de fração de ejeção neste período indicavam maior tendência de pertencerem a tal evento. A redução do pico de velocidade da onda e' da região lateral do ventrículo esquerdo, obtida pelo Doppler tecidual, reflete relaxamento ventricular mais lento nesta região, ou seja, disfunção diastólica sendo destacada pelo modelo executado, apesar da evidência ainda não comprovada na literatura. (25)

Outra variável destacada foi a TAPSE, relacionado com a função sistólica de ventrículo direito, que poderá ser afetado igualmente pelos efeitos tóxicos dos antracíclicos. (26) O comprometimento da função diastólica tem sido relacionado com a cardiotoxicidade em vários estudos, como demonstrado por metanálise, que destacou entre os parâmetros relacionado a velocidade da onda e' lateral, além de outros parâmetros sistólicos como onda S' lateral, também destacada em nosso trabalho. (27) (28)

O índice de desempenho miocárdico, que reflete a função global cardíaca sistólica e diastólica, como já descrito anteriormente em revisão sistemática, apresenta relação direta com a dosagem dos antracíclicos, e está entre os principais parâmetros ecocardiográficos relacionados com acurácia para detecção de cardiotoxicidade em diferentes estudos.

Finalmente, foi apresentado como a quarta variável, em ordem de importância, o MAPSE, que tem sido descrito como uma ferramenta capaz de detectar precocemente disfunção ventricular, inclusive sendo comparada com o SLG, mas existem poucos estudos na literatura com pacientes portadores de algum tipo de câncer. Uma redução na medida do MAPSE após tratamento foi observado, mas a implicação clínica deste achado é incerta na detecção precoce de cardiotoxicidade por antracíclicos.<sup>35</sup> Ao se observar quantidade de pacientes considerados como

evento positivo no presente estudo, que é um total de 29, tem-se que destes, cardiotoxicidade estava presente em 17 pacientes e óbitos em 12 pacientes. Dessa forma, ao se considerar o resultado apresentado pela floresta aleatória, já era esperado que valores reduzidos da FEVE fossem para pacientes em óbito ou em cardiotoxicidade clínica.

A inteligência artificial já é uma realidade em várias áreas do conhecimento, com destaque especial na área cardiovascular, com avanços na eletrocardiografia, imagem cardiovascular, algorítmicos diagnósticos, que podem mudar a história natural de muitas patologias como as consequentes dos efeitos tóxicos do tratamento do câncer.

### **5.3 Limitações do estudo**

Portanto, pode-se levantar a hipótese que existem variações além da redução da FEVE no grupo de pacientes com cardiotoxicidade, que poderia servir como preditores. Logo, para uma base de dados de pequenas dimensões como a apresentada neste trabalho, pode-se fazer boas previsões em modelos que acertam na maior parte do tempo, entretanto, não se sabe com clareza quais variáveis poderiam ser as mais relevantes como ferramenta de predição de cardiotoxicidade, principalmente pela amostra reduzida, que foi realizada em um único centro, utilizando exclusivamente o ecocardiograma bidimensional.

## 6 CONCLUSÃO

No presente estudo em pacientes com câncer de mama, submetidas ao tratamento com antracíclicos, observou-se que as seguintes variáveis ecocardiográficas como a TAPSE, o pico de velocidade da onda e' da região lateral da valva mitral, MAPSE e o índice de desempenho miocárdico , merecem atenção dos pesquisadores, oncologistas e cardiologistas na detecção precoce dos efeitos deletérios ao miocárdio, secundários aos anticíclicos, além da disfunção ventricular mensurada pela FEVE e o SLG.

## BIBLIOGRAFIA

1. Hajjar LA, Costa IBS da S da, Lopes MACQ, Hoff PMG, Diz MDPE, Fonseca SMR, et al. Diretriz Brasileira de Cardio-oncologia – 2020. *Arq Bras Cardiol* [Internet]. 2020 Nov 18;115(5):1006–43. Available from: <http://abccardiol.org/article/diretriz-brasileira-de-cardio-oncologia-2020/>
2. Lyon AR, López-Fernández T, Couch LS, Asteggiano R, Aznar MC, Bergler-Klein J, et al. 2022 ESC Guidelines on cardio-oncology developed in collaboration with the European Hematology Association (EHA), the European Society for Therapeutic Radiology and Oncology (ESTRO) and the International Cardio-Oncology Society (IC-OS). *Eur Heart J Cardiovasc Imaging*. 2022 Sep;23(10):e333–465.
3. Lyon AR, Dent S, Stanway S, Earl H, Brezden-Masley C, Cohen-Solal A, et al. Baseline cardiovascular risk assessment in cancer patients scheduled to receive cardiotoxic cancer therapies: a position statement and new risk assessment tools from the Cardio-Oncology Study Group of the Heart Failure Association of the European Society of Cardiology in collaboration with the International Cardio-Oncology Society. *Eur J Heart Fail*. 2020 Nov;22(11):1945–60.
4. Nebigil CG, Désaubry L. Updates in Anthracycline-Mediated Cardiotoxicity. *Front. Pharmacol.* 2018;9(November):1–13.
5. Herrmann J, Lenihan D, Armenian S, Barac A, Blaes A, Cardinale D, et al. Defining cardiovascular toxicities of cancer therapies: an International Cardio-Oncology Society (IC-OS) consensus statement. *Eur Heart J*. 2022 Jan 31;43(4):280–99.
6. Martín García A, Mitroi C, Mazón Ramos P, García Sanz R, Virizuela JA, Arenas M, et al. Stratification and management of cardiovascular risk in cancer patients. A consensus document of the SEC, FEC, SEOM, SEOR, SEHH, SEMG, AEEMT, AEEC, and AECC. *Rev Espanola Cardiol Engl Ed*. 2021 May;74(5):438–48.
7. Caspani F, Tralongo AC, Campiotti L, Asteggiano R, Guasti L, Squizzato A. Prevention of anthracycline-induced cardiotoxicity: a systematic review and meta-analysis. *Intern Emerg Med*. 2021 Mar;16(2):477–86.
8. Shapira J, Gotfried M, Lishner M, Ravid M. Reduced cardiotoxicity of doxorubicin by a 6-hour infusion regimen. A prospective randomized evaluation. *Cancer*. 1990 Feb 15;65(4):870–3.
9. Macedo AVS, Hajjar LA, Lyon AR, Nascimento BR, Putzu A, Rossi L, et al. Efficacy of Dexrazoxane in Preventing Anthracycline Cardiotoxicity in Breast Cancer. *JACC CardioOncology*. 2019 Sep;1(1):68–79.
10. Yamaguchi N, Fujii T, Aoi S, Kozuch PS, Hortobagyi GN, Blum RH. Comparison of cardiac events associated with liposomal doxorubicin, epirubicin and doxorubicin in breast cancer: a Bayesian network meta-analysis. *Eur J Cancer Oxf Engl* 1990. 2015

Nov;5. Nov 18;115(5):1006–43. Available from: <http://abccardiol.org/article/diretriz-brasileira-de-cardio-oncologia-2020/>

11 Oikonomou EK, Kokkinidis DG, Kampaktis PN, Amir EA, Marwick TH, Gupta D, et al. Assessment of Prognostic Value of Left Ventricular Global Longitudinal Strain for Early Prediction of Chemotherapy-Induced Cardiotoxicity: A Systematic Review and Meta-analysis. *JAMA Cardiol.* 2019 Oct 1;4(10):1007–18.

12. Melo MDT de, Paiva MG, Santos MVC, Rochitte CE, Moreira V de M, Saleh MH, et al. Brazilian Position Statement on the Use Of Multimodality Imaging in Cardio-Oncology - 2021. *Arq Bras Cardiol.* 2021 Oct;117(4):845–909.

13. Thavendiranathan P, Negishi T, Somerset E, Negishi K, Penicka M, Lemieux J, et al. Strain-Guided Management of Potentially Cardiotoxic Cancer Therapy. *J Am Coll Cardiol.* 2021 Feb;77(4):392–401.

14. Negishi T, Thavendiranathan P, Penicka M, Lemieux J, Murbraech K, Miyazaki S, et al. Cardioprotection Using Strain-Guided Management of Potentially Cardiotoxic Cancer Therapy: 3-Year Results of the SUCCOUR Trial. *JACC Cardiovasc Imaging.* 2023 Mar;16(3):269–78.

15. Melo MDT de, Salemi VMC. Usefulness of Myocardial Deformation Indices in Preventing Cardiotoxicity in Breast Cancer Patients. *Arq Bras Cardiol.* 2019 Jan;112(1):57–8.

16. Flachskampf FA, Blankstein R, Grayburn PA, Kramer CM, Kwong RYK, Marwick TH, et al. Global Longitudinal Shortening: A Positive Step Towards Reducing Confusion Surrounding Global Longitudinal Strain. *JACC Cardiovasc Imaging.* 2019 Aug;12(8 Pt 1):1566–7.

17. Quer G, Arnaout R, Henne M, Arnaout R. Machine Learning and the Future of Cardiovascular Care: JACC State-of-the-Art Review. *J Am Coll Cardiol.* 2021 Jan 26;77(3):300–13.

18. Bennett S, Cubukcu A, Wong CW, Griffith T, Oxley C, Barker D, et al. The role of the Tei index in assessing for cardiotoxicity from anthracycline chemotherapy: a systematic review. *Echo Res Pract.* 2021 May;8(1): R1–11.

19. Mitchell C, Rahko PS, Blauwet LA, Canaday B, Finstuen JA, Foster MC, et al. Guidelines for Performing a Comprehensive Transthoracic Echocardiographic Examination in Adults: *Recommendations from the American Society of Echocardiography.* *J Am Soc Echocardiogr Off Publ Am Soc Echocardiogr.* 2019 Jan;32(1):1–64.

20. Tassan-Mangina S, et al. Tissue Doppler imaging and conventional echocardiography after anthracycline treatment in adults: early and late alterations of left ventricular function during a prospective study. *Eur J Echocardiogr.* 2006;7(2):141–6.

21. Mitchell C, Rahko PS, Blauwet LA, Canaday B, Finstuen JA, Foster MC, et al. Guidelines for Performing a Comprehensive Transthoracic Echocardiographic Examination in Adults: *Recommendations from the American Society of Echocardiography*. J Am Soc Echocardiogr Off Publ Am Soc Echocardiogr. 2019 Jan;32(1):1–64.
22. Rudski LG, et al. Guidelines for the echocardiographic assessment of the right heart in adults: a report from the American Society of Echocardiography endorsed by the European Association of Echocardiography, a registered branch of the European Society of Cardiology, and the Canadian Society of Echocardiography. J Am Soc Echocardiogr. 2010;23(7):685-713.
23. Mitchell C, Rahko PS, Blauwet LA, Canaday B, Finstuen JA, Foster MC, et al. Guidelines for Performing a Comprehensive Transthoracic Echocardiographic Examination in Adults: *Recommendations from the American Society of Echocardiography*. J Am Soc Echocardiogr Off Publ Am Soc Echocardiogr. 2019 Jan;32(1):1–64.
24. Mohri M, Rostamizadeh A, Talwalkar AL A. Foundations of machine learning. MIT press, 2018.
25. Naguib M, Nixon JV, Kontos MC. Ability of Nonstrain Diastolic Parameters to Predict Doxorubicin-Induced Cardiomyopathy: A Systematic Review With Meta-Analysis. Cardiol Rev. 2018 Feb;26(1):29–34.
26. Rudski LG, et al. Guidelines for the echocardiographic assessment of the right heart in adults: a report from the American Society of Echocardiography endorsed by the European Association of Echocardiography, a registered branch of the European Society of Cardiology, and the Canadian Society of Echocardiography. J Am Soc Echocardiogr. 2010;23(7):685-713.
27. Tassan-Mangina S, et al. Tissue Doppler imaging and conventional echocardiography after anthracycline treatment in adults: early and late alterations of left ventricular function during a prospective study. Eur J Echocardiogr. 2006;7(2):141-6.
28. Støylen A, Mølmen HE, Dalen H. Relation between Mitral Annular Plane Systolic Excursion and Global longitudinal strain in normal subjects: the HUNT study. Echocardiography. 2018;35(5):603-10.
29. Hu K, et al. Clinical implication of mitral annular plane systolic excursion for patients with cardiovascular disease. Eur Heart J Cardiovasc Imaging. 2013;14(3):205-12.
30. Chen, T.; Guestrin, C. XGBoost: a Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16, p. 785–794, 2016.
31. Miyoshi T, Tanaka H, Kaneko A, et al. Left ventricular endocardial dysfunction in

patients with preserved ejection fraction after receiving anthracycline. *Echocardiography*. 2014;31:848–857.

32. Gibson DG, Francis DP. Clinical assessment of left ventricular diastolic function. *Heart*. 2003;89:231–238.

33. Arnlöv J, Ingelsson E, Risérus U, et al. Myocardial performance index, a Doppler-derived index of global left ventricular function, predicts congestive heart failure in elderly men. *Eur Heart J*. 2004;25:2220–2225.

34. Appleton CP, Hatle LK, Popp RL. Relation of transmitral flow velocity patterns to left ventricular diastolic function: new insights from a combined hemodynamic and Doppler echocardiographic study. *J Am Coll Cardiol*. 1988;12:426–440.

35. Popović ZB, Desai MY, Buakhamsri A, et al. Predictors of mitral annulus early diastolic velocity: impact of long-axis function, ventricular filling pattern, and relaxation. *Eur J Echocardiogr*. 2011; 12:818–825.

36. Kulkarni P, Mahadevappa M, Chilakamarri S. The Emergence of Artificial Intelligence in Cardiology: Current and Future Applications. *Curr Cardiol Rev*. 2022;18(3):e191121198124.