



## Uso de machine learning no diagnóstico de câncer de mama através de ultrassonografia

### Using machine learning in breast cancer diagnosis through ultrasonography

### Uso de machine learning en el diagnóstico de cáncer de mama a través de ultrasonografía

Marcelo Rosano Dallagassa<sup>1</sup>, Antonio Josenias Cordeiro de Oliveira<sup>2</sup>

1 Doutorado/especialista, NIIS, Unimed, Curitiba, PR, Brasil

2 Especialista em Big Data, Gestão de Dados e Analytics, Politécnico, PUC PR, Curitiba, PR, Brasil

Autor correspondente: Antonio Josenias Cordeiro de Oliveira

E-mail: joseniascordeiro95@gmail.com

#### Resumo

Objetivo: Aplicar técnicas de *machine learning* para avaliar a eficiência e precisão no diagnóstico de câncer de mama por meio de imagens de ultrassonografia. Métodos: Utilizando a ferramenta Orange Data Mining, foram processadas imagens de um banco obtido na plataforma Kaggle contendo ultrassonografias de mamas que apresentavam nódulos benignos e malignos. Resultados: A análise das imagens nos testes foi realizada por meio das técnicas de *machine learning*, *Neural Network* e KNN, que demonstraram altas taxas de precisão na classificação de nódulos benignos e malignos com resultados acima de 90% de assertividade. Conclusão: A pesquisa sublinha a importância de ferramentas de *machine learning* para criar *modelos* preditivos robustos e precisos, melhorando significativamente a precisão do diagnóstico clínico e aumentando as chances de recuperação dos pacientes.

**Descritores:** Câncer da Mama; Aprendizado de Máquina; Ultrassonografia Mamária

#### Abstract

Objective: To apply machine learning techniques to evaluate the efficiency and accuracy in diagnosing breast cancer using ultrasound images. Methods: Using



the Orange Data Mining tool, images from a dataset obtained on the Kaggle platform containing ultrasound scans of breasts with benign and malignant nodules were processed. Results: Image analysis in the tests was performed using machine learning techniques, Neural Network and KNN, which demonstrated high accuracy rates in classifying benign and malignant nodules with results above 90% accuracy. Conclusion: The research emphasizes the importance of machine learning tools to create robust and accurate predictive models, significantly improving clinical diagnostic accuracy and increasing patient recovery chances.

**Keywords:** Breast Cancer; Machine Learning; Ultrasonography, Mammary

## Resumen

Objetivo: Aplicar técnicas de machine learning para evaluar la eficiencia y precisión en el diagnóstico de cáncer de mama mediante imágenes de ultrasonido. Métodos: Utilizando la herramienta Orange Data Mining, se procesaron imágenes de un banco de datos obtenido en la plataforma Kaggle que contenía ultrasonidos de mamas con nódulos benignos y malignos. Resultados: El análisis de las imágenes en las pruebas se realizó mediante técnicas de machine learning, Neural Network y KNN, que demostraron altas tasas de precisión en la clasificación de nódulos benignos y malignos con resultados por encima del 90% de precisión. Conclusión: La investigación subraya la importancia de las herramientas de machine learning para crear modelos predictivos robustos y precisos, mejorando significativamente la precisión del diagnóstico clínico y aumentando las posibilidades de recuperación de los pacientes.

**Descriptores:** Cáncer de Mama; Aprendizaje Automático; Ultrasonografía Mamaria

## Introdução

O conceito de machine learning (ML) é definido como um ramo da inteligência artificial (IA) que se dedica ao desenvolvimento de sistemas capazes de aprender e aprimorar seu desempenho a partir dos dados que recebem <sup>(1)</sup>. A *machine learning* é



uma parte crucial do crescente campo da ciência de dados. Através de métodos estatísticos, os algoritmos são treinados para realizar classificações ou previsões, proporcionando insights essenciais em projetos de mineração de dados <sup>(2)</sup>.

Nas últimas décadas, os avanços tecnológicos têm possibilitado a criação de vários produtos baseados em *machine learning*, inclusive no ramo da medicina. Estas tecnologias passaram a ser fundamentais no papel do cuidado à saúde da sociedade, sendo amplamente utilizadas nos diversos ramos da medicina. O machine learning encontra uma de suas principais aplicações no prognóstico e predição do câncer <sup>(3)</sup>.

O câncer, também conhecido como neoplasia, é um conjunto de mais de 100 doenças que possuem como característica comum o crescimento atípico de células (. Entre essas doenças, o câncer de mama se destaca como o tipo mais comum entre mulheres no Brasil e no mundo, conforme dados do Instituto Nacional de Câncer (INCA). Segundo o Ministério da Saúde, o câncer de mama representa cerca de 28% dos novos casos de câncer em mulheres no país <sup>(4,5)</sup>.

Em 2020 a Organização Mundial da Saúde (OMS) relatou que 2,3 milhões de mulheres foram diagnosticadas com câncer de mama, resultando em 658 mil mortes globalmente, destacando a alta prevalência mundial da doença. Conforme dados do INCA estimou-se que o Brasil teria 73.610 novos casos de câncer de mama em 2023, com um risco estimado de 66,54 casos por 100 mil mulheres. Além disso, o câncer de mama é a principal causa de mortalidade por câncer entre mulheres no país <sup>(6, 7)</sup>.

O principal sintoma deste tipo de câncer trata-se do surgimento de nódulos nos seios. Esses nódulos, no entanto, podem ser classificados como benignos ou malignos, ou seja, podem ou não serem cancerígenos. Porém, esse não é seu único sintoma. O Ministério da Saúde também destaca outros sintomas da doença, incluindo retração da pele, secreção papilar (especialmente quando unilateral e espontânea), dor, inversão do mamilo, hiperemia, edema cutâneo (semelhante à casca de laranja) e descamação ou ulceração do mamilo <sup>(5)</sup>.



Um estudo realizado por McKinsey <sup>(8)</sup>, aponta que A Inteligência Artificial mostra grande capacidade para identificar o câncer de mama. Em seu artigo, o autor destaca dados que indicam elevadas incidências de resultados falso positivos e falso negativos nas mamografias interpretadas por radiologistas. No entanto, essa vulnerabilidade é significativamente reduzida quando o mapeamento é realizado com o auxílio da IA. A pesquisa realizada por Dembrower <sup>(9)</sup> reforça essa afirmativa. Em seu artigo os autores chegaram à conclusão, através de testes e análises, que o uso de IA para triagem de mamografias sem avaliação radiológica e avaliação aprimorada poderia reduzir potencialmente a carga de trabalho do radiologista em mais da metade e detectar preventivamente uma proporção substancial de cânceres diagnosticados posteriormente.

A mamografia é a principal ferramenta para diagnóstico de câncer de mama no Brasil. Contudo, mesmo com a disponibilidade gratuita do exame, as taxas de novos casos e de mortalidade seguem elevadas. Esse fenômeno ocorre em parte porque a cobertura do exame mamográfico não é homogênea em todo o país e fatores socioeconômicos também afetam o acesso ao exame <sup>(10)</sup>.

Além disso, segundo Aguiar <sup>(11)</sup>, como essa técnica utiliza a emissão de feixes de Raios X para gerar imagens da mama, sua precisão pode ser prejudicada em pacientes jovens, pois elas possuem mamas mais densas, o que dificulta a geração de contraste pelos Raios X. Neste cenário, a ultrassonografia das mamas desempenha um papel crucial como exame complementar à mamografia e avaliação clínica, estabelecendo-se como um exame de imagem fundamental para o diagnóstico de doenças mamárias <sup>(12)</sup>.

A ultrassonografia apresenta diversas vantagens, sendo um método rápido, acessível, de custo reduzido e que não requer compressão mamária nem expõe a paciente à radiação ionizante. A inteligência artificial é extensivamente empregada na análise de imagens e em previsões. Várias especialidades médicas têm trabalhado no desenvolvimento de técnicas para refinar a IA, com o objetivo de promover avanços na saúde de forma amplamente reconhecida <sup>(13, 14)</sup>.

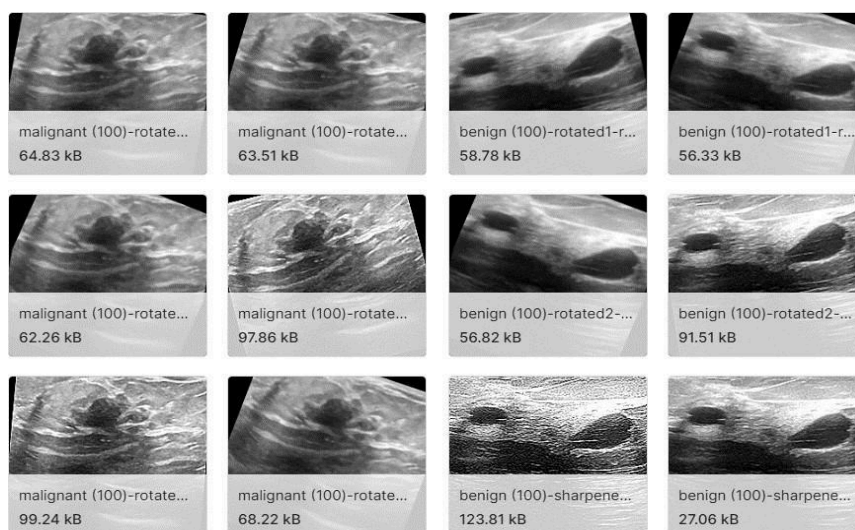


Nesse sentido, o desenvolvimento de um algoritmo capaz de auxiliar a realização do diagnóstico de tumores através da análise de exames imagens pode auxiliar médicos e pacientes a terem uma resposta mais rápida, considerando que o diagnóstico precoce aumenta a probabilidade de sucesso no tratamento de cura. Neste cenário, este projeto tem como objetivo utilizar o exame de ultrassonografia em conjunto com o machine learning no diagnóstico precoce do câncer de mama, destacando os benefícios dessa abordagem integrada na prática clínica e suas implicações para o tratamento da doença.

## Metodologia

A primeira fase do experimento consistiu na coleta dos dados que seriam usados nos testes. Para este experimento, foi utilizado um banco de imagens obtido através da plataforma Kaggle <sup>(15)</sup>. O conjunto de dados em questão reunia apenas imagens de ultrassonografias mamárias, cujos exames apontaram a existência de nódulos, que por sua vez foram classificados como sendo benignos ou malignos, conforme ilustrado na figura 1.

**Figura 1** - Ultrassonografias mamárias



Fonte: Kaggle

A ferramenta utilizada para o treinamento dos dados foi o Orange Data Mining, um *software* de código aberto usada para o treinamento de dados, destacando-se pela simplicidade de uso, já que não exige codificação. É capaz de processar linguagem natural e mineração de texto, além de realizar mineração de regras de associação e análises de rede. A ferramenta oferece diversos elementos estatísticos, como gráficos, árvores de decisão, mapas de calor e projeções lineares, e permite a instalação de bibliotecas adicionais para expandir suas funcionalidades.

Após realizada a coleta dos dados que seriam utilizados no treinamento, foi necessário preparar o ambiente para recebê-los. Nessa etapa, foi feita a instalação da biblioteca *Image Analytics*, que não compunha os pacotes de *widgets* previamente instalados no Orange, necessária para viabilizar o *upload* de imagens, e consequentemente possibilitando o trabalho com esse tipo de dado. Para que seja possível alcançar o objetivo deste projeto a utilização desta biblioteca é indispensável.

No Orange todo o fluxo lógico é estruturado através de *widgets* arrastáveis, que se interligam entre si. Para os experimentos deste projeto foram criados dois fluxos, um para treinamento e o outro para os testes, ambos trabalhando com o mesmo tipo de dados, imagens de ultrassonografia contendo nódulos malignos e benignos. O objetivo da criação dos dois fluxos foi utilizar um para ensinar a ferramenta a fazer a distinção correta das imagens, sendo capaz de apontar os nódulos malignos ou não, e o segundo fluxo seria responsável por colocar em prática o que foi aprendido nas análises realizadas no treinamento.

No treinamento, processamos um total de 8.116 imagens de ultrassonografia, divididas em 4.074 imagens com nódulos benignos e 4.042 imagens com nódulos malignos. Para a fase de testes, utilizamos 900 imagens ultrassonográficas, sendo 500 delas com nódulos benignos e 400 com nódulos malignos.

Para este trabalho aplicou-se o desenho de pesquisa: primário – observacional – transversal – retrospectivo – estudo de acurácia. Primário, pois é uma investigação original. Observacional, pois não interfere ou modifica qualquer aspecto que esteja



estudando. Transversal ou seccional, refere-se a um estudo que a exposição ao fator causa está presente ao efeito no mesmo momento ou intervalo de tempo analisado. Retrospectivo, o estudo considera registros do passado com a importância sobre os dados coletados para serem computados. E por fim, estudo de acurácia, pois o foco do estudo é de analisar a acurácia de modelos para teste diagnóstico <sup>(16)</sup>.

## Resultados e Discussões

Entre os elementos principais do fluxo analítico destaca-se o papel dos modelos de predição, no Orange esse grupo de elementos fazem parte de um pacote de *widgets* denominados *Models*. Os *widgets* desse segmento são responsáveis pelo processamento dos dados.

Na primeira versão criada, para o fluxo de treinamentos, foram realizados testes com 5 modelos de predição, sendo eles: *AdaBoost*, *Logistic Regression*, *KNN*, *Neural Network* e *SVM*. O objetivo da análise com os modelos citados foi identificar qual(is) modelo(s) traria(m) melhor(es) resultado(s) no processamento das imagens. De acordo com os testes de comparação realizados no *widget Test and score* foi possível identificar que a *Neural Network* e o *KNN* apresentaram melhor acurácia e precisão em relação aos demais. O *widget Test and score* faz parte do módulo do Orange denominado *Evaluate*, e é responsável por testar algoritmos de aprendizagem em dados. O *Test and score* mostra uma tabela com diferentes medidas de desempenho, como precisão de classificação e área sob a curva. Além disso, ele gera resultados que podem ser utilizados por outros *widgets* para analisar o desempenho de classificadores. O resultado da comparação pode ser observado na figura 2.

**Figura 2** - Resultado dos Modelos no Test and score

Model	AUC	CA	F1	Prec	Recall	MCC
Neural Network	0.997	0.972	0.972	0.972	0.972	0.944
SVM	0.827	0.743	0.742	0.748	0.743	0.491
kNN	0.988	0.950	0.950	0.950	0.950	0.900
Logistic Regression	0.983	0.931	0.931	0.931	0.931	0.862
AdaBoost	0.829	0.829	0.829	0.829	0.829	0.657

Visto que, a *Neural Network* e o *KNN* apresentaram os melhores resultados, optou-se por seguir apenas com os dois modelos. A decisão do não seguimento com os demais modelos é justificada pelo ganho de performance, a utilização de muitos elementos deste tipo resultaria em um maior tempo para o processamento dos dados.

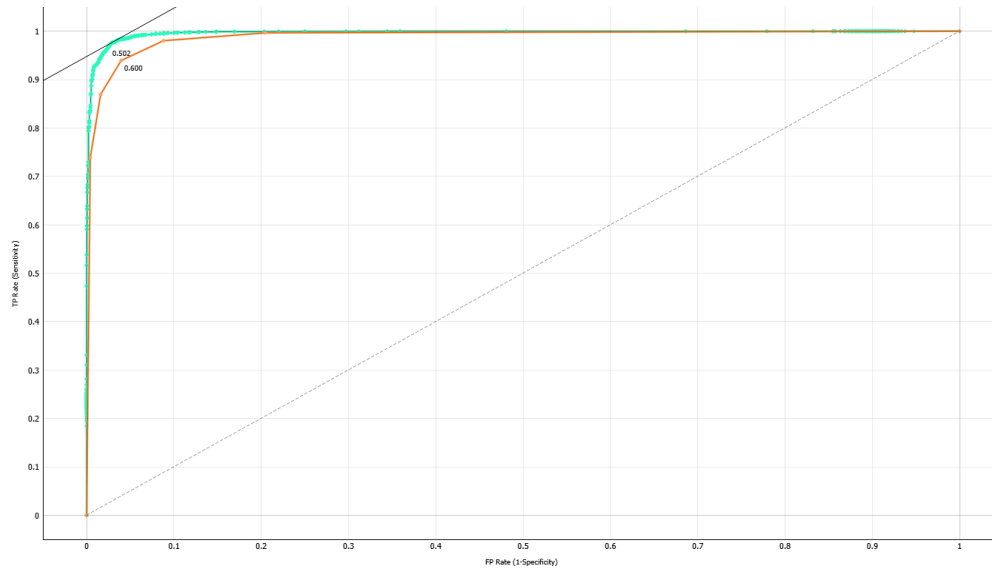
O *Neural Network* simula o processo decisório do cérebro humano, replicando a atividade dos neurônios biológicos. Já o *KNN* identifica, em um conjunto específico, os "K" elementos mais semelhantes ao elemento a ser classificado, baseado na suposição de que amostras mais semelhantes estarão mais próximas no espaço multidimensional definido pelas variáveis.

Foi adicionado ao fluxo de treinamento um *widget* chamado *ROC Analysis* para comparar graficamente os modelos de classificação. A análise *ROC* mostra que quanto mais próxima a curva estiver da borda esquerda e superior do espaço *ROC*, mais preciso é o classificador. Através do *ROC Analysis*, observou-se que a *Neural Network* teve uma performance levemente superior ao *KNN*, tanto para imagens de nódulos benignos quanto para imagens com câncer. Os resultados da comparação podem ser observados nas figuras 3 e 4.

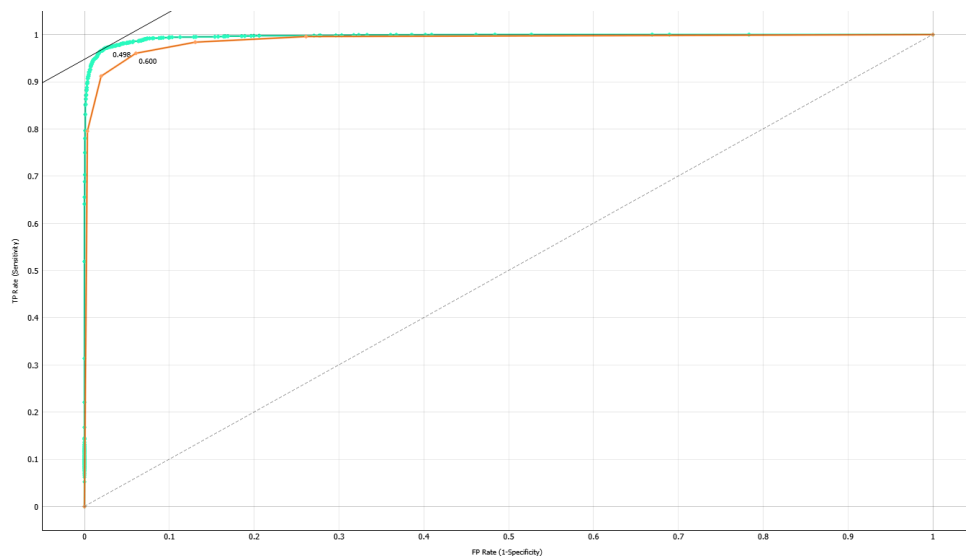




**Figura 3 - ROC Analysis com ultrassonografia de nódulos benignos**



**Figura 4 - ROC Analysis com ultrassonografia de nódulos malignos**



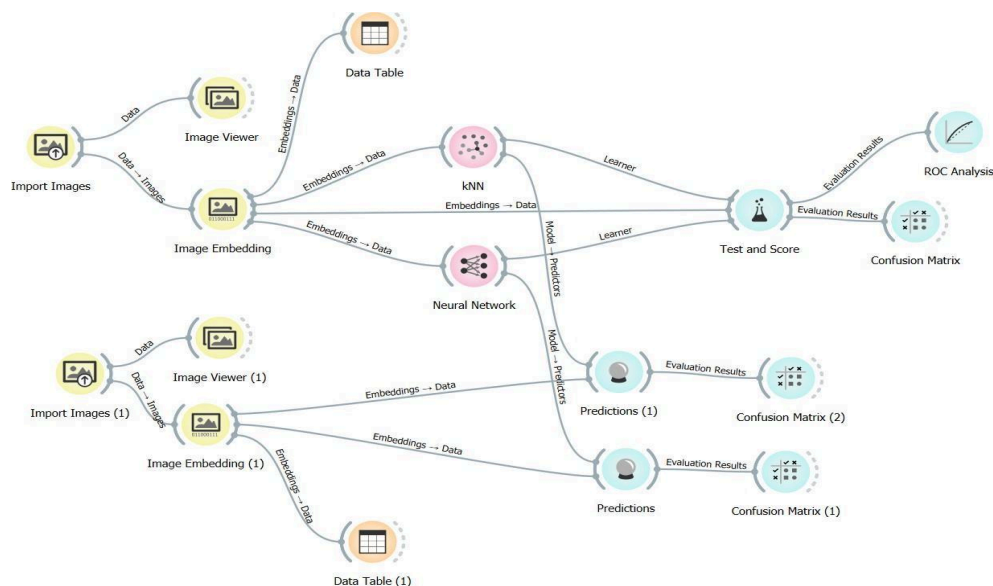
No fluxo de testes, diferentemente do fluxo de treinamento, foram inseridos dois avaliadores do tipo *Predictions*. Ele recebe um conjunto de dados de um ou mais modelos preditivos e gera dados de previsões. O *Predictions* é capaz de mostrar as probabilidades e decisões dos modelos preditivos. O *Predictions* e o *ROC Analysis* pertencem ao mesmo módulo ao qual o *Test and score* faz parte.



Cada um dos *Predictions* recebeu dados de um modelo preditivo diferente, um para o KNN e outro para a *Neural Network*, permitindo uma comparação mais clara dos resultados ao final dos testes. Os *Predictions* conectaram os modelos preditivos no fluxo de treinamento com o fluxo de testes, sendo responsáveis por mensurar o nível de sucesso do experimento.

O esquema geral criado no *workflow* do Orange para analisar as imagens de ultrassonografia pode ser observado na figura 5.

**Figura 5** - Esquema final construído no Orange



O resultado do treinamento com os dois modelos utilizados pôde ser observado por meio de uma *Confusion Matrix*, um *widget* que é geralmente utilizado para obter os resultados da avaliação de um *Test and score*. A *Confusion Matrix* tem como função fornecer os números e/ou proporções de assertividade entre as classes.

Os resultados do treinamento apresentados pela *Confusion Matrix* mostraram que a *Neural Network* classificou corretamente 3.952 das 4.074 imagens de nódulos benignos, errando um total de 122 imagens, ou seja, obtendo uma taxa de 97% de assertividade. Em relação às imagens com câncer de mama, a *Neural Network* classificou corretamente 3.937 do total de 4.042 imagens, errando somente 105. Um



total de 97,4% de assertividade. O resultado dos dados obtidos com os treinamentos utilizando a *Neural Network* pode ser observado na figura 6.

**Figura 6** - Resultado dos treinamentos obtidos através do Neural Network

		Predicted		$\Sigma$
		benign	malignant	
Actual	benign	3952	122	4.074
	malignant	105	3937	4.042
$\Sigma$		4.057	4.059	8.116

Os resultados dos treinamentos dos dados utilizando o KNN tiveram um desempenho abaixo em relação ao *Neural Network*, porém também apresentaram números satisfatórios de assertividade. O KNN acertou 3.827 e errou 247 das imagens de nódulos benignos, obtendo um total de 93,9% de acertos. E acertou 3.882 imagens de nódulos malignos, errando apenas 160. Um total de 96% de assertividade. Os dados obtidos por meio da *Confusion Matrix* utilizando o KNN podem ser observados na figura 7.

**Figura 7** - Resultado dos treinamentos obtidos através do KNN

		Predicted		$\Sigma$
		benign	malignant	
Actual	benign	3827	247	4.074
	malignant	160	3882	4.042
$\Sigma$		3.987	4.129	8.116

Após verificados os resultados dos treinamentos, foi observado o nível de eficiência dos modelos de predição no fluxo de testes. A *Neural Network* previu corretamente 484 e errou 16 das imagens contendo nódulos benignos, um total de 96,8% de aproveitamento. Já em relação às imagens contendo nódulos malignos, a



*Neural Network* conseguiu acertar 372 das 400 imagens, errando um total de 28 imagens, um aproveitamento de 93%. Em números absolutos, pode-se afirmar que os resultados dos treinamentos e dos testes foram tecnicamente os mesmos em assertividade. O resultado dos testes com a *Neural Network* pode ser observado na figura 8.

**Figura 8** - Resultado dos testes obtidos através do Neural Network

		Predicted		$\Sigma$
		benign	malignant	
Actual	benign	484	16	500
	malignant	28	372	400
$\Sigma$		512	388	900

O KNN acertou 430 das imagens com nódulos benignos, errando 70, um aproveitamento de 86%. E classificou corretamente 368 imagens com nódulos malignos, errando 32 imagens, um total de 92% de assertividade. Os resultados podem ser observados na figura 9.

**Figura 9** - Resultado dos testes obtidos através do KNN

		Predicted		$\Sigma$
		benign	malignant	
Actual	benign	430	70	500
	malignant	32	368	400
$\Sigma$		462	438	900

Assim como o esperado, o KNN obteve desempenho menor em relação a *Neural Network*, também houve queda de rendimento em relação aos seus próprios números nos treinamentos, levando à conclusão que a *Neural Network* é o modelo mais indicado para esse tipo de análise. Porém, não justifica um descarte do KNN



para este e para estudos futuros, dados os bons números obtidos nos testes realizados.

Existem diversos projetos na literatura que utilizam diferentes modelos para auxiliar na predição do câncer de mama. De Rezende e Maimone <sup>(17)</sup> utilizaram técnicas de aprendizado de máquina na oncologia, especificamente a Árvore de Decisão, que, assim como a Rede Neural e o KNN, é um algoritmo de análise de dados. Em seu estudo, analisaram variáveis clínicas para identificar possíveis relações e influências sobre o número de óbitos em uma base de dados. A pesquisa concluiu que o uso de técnicas de IA, como o aprendizado de máquina e o algoritmo de Árvore de Decisão, pode fornecer insights valiosos na análise de fatores clínicos relacionados ao câncer de mama.

Diniz et al. <sup>(18)</sup> propuseram um método para diagnóstico de câncer de mama em imagens de ultrassom, combinando pré-processamento, aumento de dados balanceados e um Ensemble de EfficientNet para melhorar a precisão do diagnóstico. O método, que utiliza técnicas de processamento de imagens e aprendizado profundo, mostrou-se eficaz na melhoria da precisão do diagnóstico de câncer de mama em imagens de ultrassom.

Aguiar <sup>(11)</sup> seguiu uma linha de pesquisa semelhante à proposta neste projeto, utilizando a plataforma Orange para a classificação de imagens termográficas mamárias de pacientes do Hospital das Clínicas da UFPE. A plataforma demonstrou eficiência na classificação binária entre "Câncer e não-câncer", alcançando mais de 96% de acurácia e mais de 99% de sensibilidade ao maligno. Esses resultados mostram o potencial da Orange Canvas para futuros usuários na análise e diagnóstico médico.

A aplicação de técnicas de machine learning permite a análise eficiente de grandes volumes de dados, a identificação de padrões complexos e a melhoria na precisão diagnóstica, essenciais para diagnósticos precoces e precisos. A integração com a plataforma Orange Canvas no diagnóstico do câncer de mama por meio de



ultrassonografia representa um avanço significativo, proporcionando diagnósticos mais precisos, rápidos e acessíveis, melhorando os resultados do tratamento e a sobrevivência dos pacientes, conforme observado nas pesquisas citadas.

## Conclusão

Este estudo destaca a relevância e a eficácia das técnicas de machine learning na análise de imagens de ultrassonografia para o diagnóstico do câncer de mama. A aplicação dos modelos preditivos Neural Network e KNN alcançou altos níveis de precisão na distinção entre nódulos benignos e malignos. A Neural Network demonstrou uma capacidade superior de identificar corretamente a maioria dos casos, consolidando-se como a abordagem mais eficiente. O desempenho do KNN, embora ligeiramente inferior, também apresentou resultados satisfatórios, com taxas de acerto realmente relevantes na distinção das imagens.

Os achados sublinham a importância do uso de ferramentas de mineração de dados e machine learning, como o Orange Data Mining, na criação de modelos preditivos robustos e precisos. A capacidade de processar grandes conjuntos de dados e a flexibilidade dos widgets do Orange permitiram um fluxo de trabalho eficiente, possibilitando uma análise detalhada e comparativa dos modelos.

Em resumo, a pesquisa confirma que a integração de tecnologias avançadas de aprendizado de máquina na prática clínica pode melhorar significativamente a precisão do diagnóstico do câncer de mama. Esses avanços aumentam as chances de recuperação dos pacientes e destacam a importância do desenvolvimento contínuo e da aplicação de tecnologias inovadoras na medicina, prometendo um futuro promissor na luta contra o câncer de mama através de diagnósticos precoces e precisos.

## Referências

1. Oracle Brasil. O que é Machine learning? [202-]. Available from: <https://www.oracle.com/br/artificial-intelligence/machine-learning/what-is-machine-learning/>. Accessed 2024 Feb 17.



2. IBM. O que é machine learning? [202-]. Available from: <https://www.ibm.com/br-pt/topics/machine-learning>. Accessed 2024 Feb 18.
3. Silva ME, et al. Proposta e avaliação de um modelo híbrido de seleção de características para o prognóstico do câncer de mama. 2022.
4. O que é câncer? [Internet]. Instituto Nacional de Câncer - INCA. [citado 18 de janeiro de 2024]. Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/o-que-e-cancer>
5. Ministério da Saúde. Câncer de mama. [202-]. Available from: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/c/cancer-de-mama>. Accessed 2024 Jan 18.
6. Ferreira CSC. Uso de inteligência artificial em deep learning na segmentação da mama em imagens de ressonância magnética como ferramenta para o diagnóstico de câncer de mama: uma revisão sistemática. 2023.
7. Instituto Nacional de Câncer. Outubro Rosa 2023. (2023). Available from: <https://www.gov.br/inca/pt-br/assuntos/campanhas/2023/outubro-rosa>. Accessed 2024 Feb 19.
8. McKinney SM, et al. International evaluation of an AI system for breast cancer screening. *Nature*. 2020;577(7788):89-94.
9. Dembrower K, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. *The Lancet Digital Health*. 2020;2(9)
10. Campos KFA. Importância da mamografia no rastreamento do câncer de mama: uma revisão de literatura. 2023.
11. Aguiar YM, et al. Desenvolvimento e aplicação de metodologia de aprendizagem de máquina para classificação de imagens termográficas na área médica. 2021. Dissertação de Mestrado. Universidade Federal de Pernambuco.
12. Calas MJG, et al. Estudo preliminar das limitações técnicas da ultrassonografia automatizada da mama: do procedimento ao diagnóstico. *Radiologia Brasileira*. 2020;53:293-300.
13. Vieira WL, et al. A importância da Ultrassonografia no Câncer de Mama. Março de 2018 24ª EDIÇÃO, p. 41, 2018.
14. Braga AV, et al. Machine learning: O Uso da Inteligência Artificial na Medicina. *Brazilian Journal of Development*. 2019;5(9):16407-13.
15. Sairam VA. Ultrasound Breast Images for Breast Cancer. (2024). Available from: <https://www.kaggle.com/datasets/vuppalaadithyasairam/ultrasound-breast-images-for-breast-cancer/data>. Accessed 2024 Jan 14.
16. Hochman B, Nahas FX, Oliveira Filho RS, Ferreira LM. Desenhos de pesquisa. *Acta Cir Bras*. 2005;20(suppl 2):2-9. doi: 10.1590/S0102-86502005000800002.
17. De Rezende MR, Maimone TS. Uso de inteligência artificial para identificação de fatores influenciadores do câncer de mama a partir de dados clínicos. 2022.
18. Diniz JOB, Dias DA Jr, Cruz LB da, Marques RCS, Gomes DL Jr, Cortês OAC, et al. EfficientEnsemble: Diagnóstico de câncer de mama em imagens de ultrassom utilizando processamento de imagens e Ensemble de EfficientNets. Em: Anais do XXIV Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2024). Sociedade Brasileira de Computação - SBC; 2024. p. 202-13.