Título:	**MODELO DE RNA MULTICAMADAS PARA CLASSIFICAÇÃO DE TUMORES MAMÁRIOS UTILIZANDO O *DIAGNOSTIC WISCONSIN BREAST CANCER DATABASE***		
Autores:	Paula Camilly da Silva Daniela Duarte da Silva Bagatini João Carlos Furtado		
Área	[ ] Humanas [ ] Sociais Aplicadas [ ] Biológicas e da Saúde [x] Exatas, da Terra e Engenharias	Dimensão:	[x] Ensino [ ] Pesquisa [ ] Extensão [ ] Inovação

Resumo:

Introdução: No Brasil, o câncer de mama é o tipo mais frequente entre as mulheres, com taxa ajustada de incidência de 41,89 casos por 100 mil mulheres (INCA, 2025). Sua etiologia multifatorial torna o diagnóstico precoce determinante para a sobrevida. Nesse cenário, técnicas de Redes Neurais Artificiais (RNA), consolidam-se como recursos auxiliares capazes de identificar padrões complexos em dados clínicos. **Objetivo:** Desenvolver e avaliar uma RNA do tipo Multilayer Perceptron (MLP) para classificar tumores mamários em benignos e malignos, com ênfase na detecção de casos malignos e na redução de falsos negativos, erro que pode retardar o diagnóstico e o tratamento. **Metodologia:** Utilizou-se o dataset público Diagnostic Wisconsin Breast Cancer (569 amostras, 30 variáveis; 357 benignas e 212 malignas), com características de núcleos celulares extraídas de punção aspirativa por agulha fina. O pré-processamento incluiu verificação de valores nulos, remoção da coluna identificadora, recodificação da variável diagnóstica, normalização e divisão estratificada em 455 amostras de treino (285 benignas, 170 malignas) e 114 de teste (72 benignas, 42 malignas). A rede e a interface gráfica (Tkinter) foram desenvolvidas com TensorFlow/Keras. Aplicou-se Grid Search com duas arquiteturas de MLP (16 e 32 neurônios na 1ª camada, 8 e 16 na 2ª), dois valores de dropout (0,2 e 0,3) e duas taxas de aprendizagem (0,001 e 0,0005), totalizando 16 combinações avaliadas pela acurácia no teste. O modelo final incorporou EarlyStopping (paciência de 10 épocas na perda de validação). **Resultados:** A melhor configuração (16-8 neurônios, dropout 0,2, taxa 0,001) convergiu em 37 épocas e alcançou acurácia de 99,12% nas 114 amostras de teste. A matriz de confusão registrou 72 verdadeiros negativos, 0 falso positivo, 1 falso negativo e 41 verdadeiros positivos. Por classe, obteve-se, para benignos, precisão 0,9863, recall 1,000 e F1-Score 0,9931; para malignos, precisão 1,000, recall 0,9762 e F1-Score 0,9880. As médias macro foram 0,9932, 0,9881 e 0,9905, e as ponderadas, 0,9913, 0,9912 e 0,9912, indicando equilíbrio entre as classes diante do desbalanceamento do dataset (63,2% benignas vs. 36,8% malignas). A convergência das curvas de treino e validação evidenciou boa generalização. A análise de importância apontou maior influência de atributos morfológicos nucleares. Dois testes com exemplos predefinidos confirmaram a coerência do modelo (100% de probabilidade de malignidade no caso maligno e 2,11% no benigno). **Conclusão:** O modelo MLP constitui alternativa promissora como ferramenta de apoio ao diagnóstico do câncer de mama e contribui à redução de falsos negativos e para o ODS 3 (Saúde e Bem-Estar) ao ampliar o

acesso a triagens rápidas. Como limitações, destacam-se o tamanho reduzido e a origem única do dataset, que restringem a generalização; a avaliação em partição única, tornando as métricas sensíveis à divisão adotada; a seleção do melhor modelo pela acurácia no próprio conjunto de teste, com risco de otimismo; o uso de medidas extraídas por especialistas, e não de imagens brutas; e a estimativa de importância baseada nos pesos da primeira camada. Como perspectivas futuras, destacam-se a validação cruzada e externa, o ajuste do limiar de decisão voltado à classe maligna e a incorporação de interpretabilidade.

Instituto Nacional de Câncer. Controle do câncer de mama no Brasil: dados e números 2025/Instituto Nacional de Câncer. Rio de Janeiro: INCA, 2025.

Link do Vídeo:
https://drive.google.com/drive/folders/1jT7H_Izn2RInGu8ii1KKoIcnmW7pjvhw.