

Conheça o Problema

O câncer de mama é uma doença em que células anormais se multiplicam descontroladamente na mama, formando tumores malignos ou benignos. E é o câncer mais comum e letal entre mulheres do mundo todo.

Pensando nesses dados desenvolvemos um modelo de machine learning para classificar os tumores de mama como malignos ou benignos.



DataSet

Breast Cancer Wisconsin (Diagnostic)

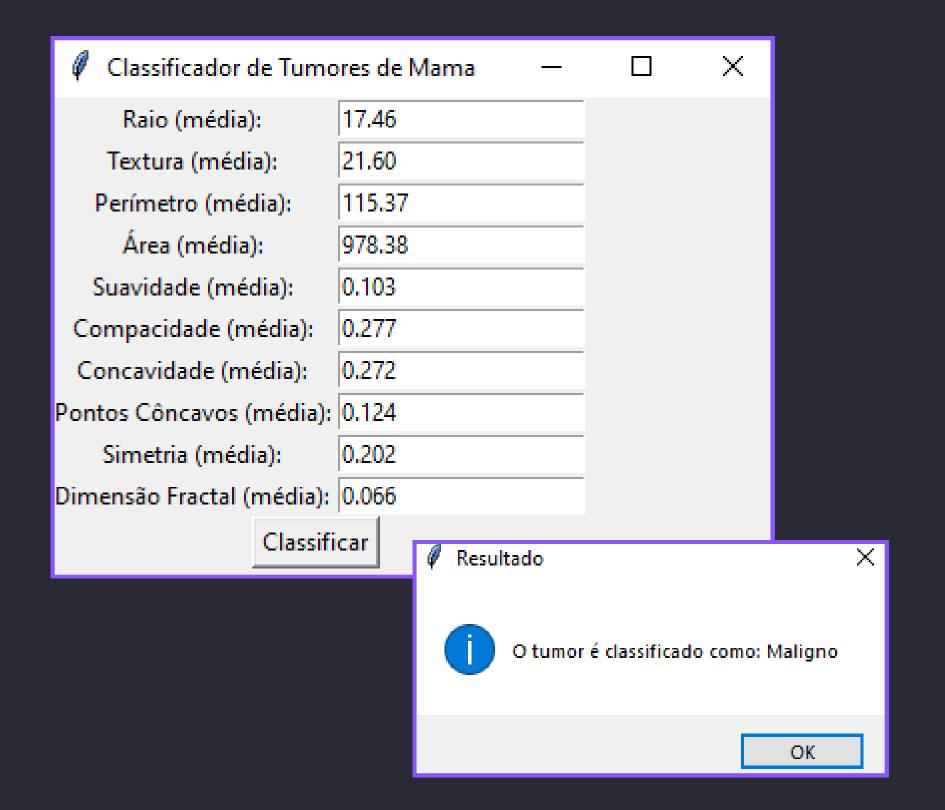
• O DataSet Breast Cancer Wisconsin (Diagnostic) conta com 569 casos sendo 212 desses casos malignos e 357 benignos

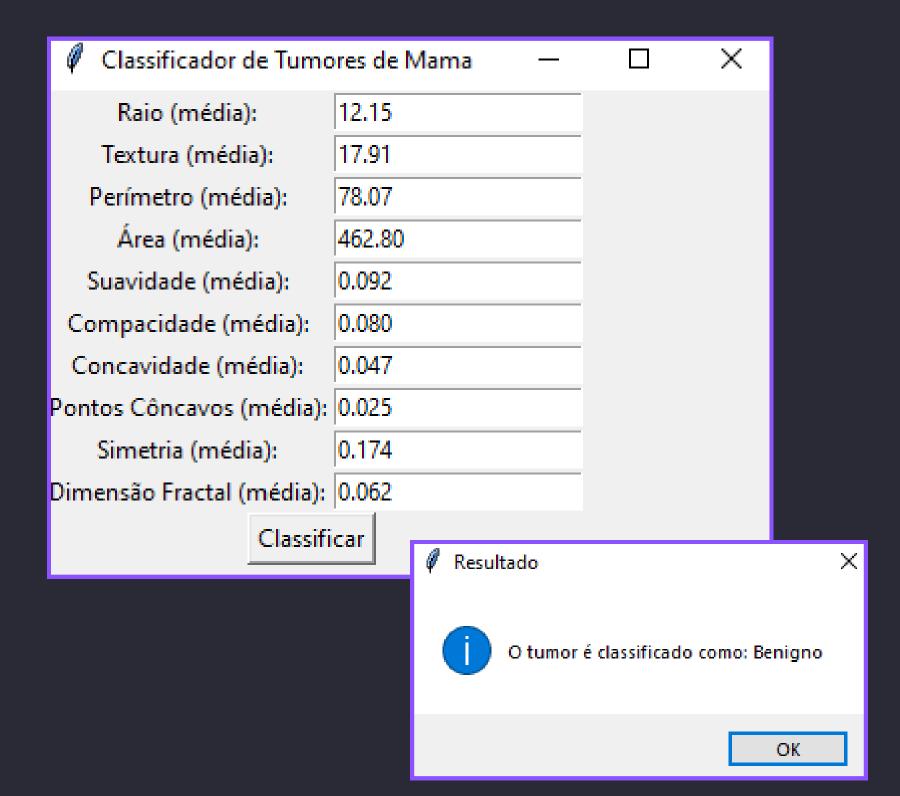
Ele é frequentemente utilizado em pesquisas e competições de machine learning devido à sua qualidade e relevância clínica

DataSet

Breast Cancer Wisconsin (Diagnostic)

- O objetivo do DataSet é auxiliar no diagnóstico de tumores de mama, classificando como malignos ou benignos com base em características extraídas de imagens de tecido mamário. O dataset é altamente relevante para a área de saúde, pois pode auxiliar médicos no diagnóstico precoce e preciso de câncer de mama
- O DataSet Utiliza 30 características numéricas, calculadas a partir de imagens digitalizadas de núcleos celulares.
 Mas neste caso foi utilizada apenas as 10 primeiras características como base





Acurácia do modelo k-NN: 94.15%

Acurácia média com validação cruzada: 93.85%

Desvio padrão das acurácias: 0.97%

Conclusão

- Em resumo o modelo alcançou uma Eficácia de 94,15%, demonstrando ser muito eficaz na classificação dos tumores.
- Existem diversos projetos que utilizam machine learning para dar assistencia a medicos em diagnosticos do mundo todo, um deles seria o Google DeepMind realizado em parceria com o Serviço Nacional de Saúde do Reino Unido que ja reduziu em 9,4% os falsos negativos.



