

Renan Maeda de Souza; Luis Miguel Leuch; Arthur Garani Narciso; Marcelo de Oliveira Dreweck (Universidade Estadual de Ponta Grossa). Autor correspondente: Marcelo Oliveira Dreweck, modreweck@uepg.com.br

Introdução

O acúmulo de gordura hepática está associado a doenças crônicas hepáticas e cardiovasculares, exigindo diagnósticos precoces e eficazes (Hu et al., 2023). A tomografia computadorizada (TC) permite análise detalhada do fígado, enquanto a radiômica extrai dados quantitativos de imagens, como textura e intensidade, contribuindo para modelos preditivos não invasivos (Sack et al., 2022). No entanto, a alta correlação entre variáveis radiômicas pode causar multicolinearidade, comprometendo a estabilidade e interpretação dos modelos (Wang et al., 2020; Tang et al., 2023). Estratégias estatísticas, como regressões RIDGE e LASSO, ajudam a mitigar esse problema, melhorando a precisão e robustez dos diagnósticos (Hu et al., 2023; Wang et al., 2022). Este estudo propõe comparar essas abordagens na acurácia de modelos radiômicos para quantificação da gordura hepática pela TC.

Objetivo

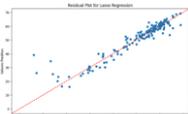
Comparar diferentes estratégias de pré-processamento de dados que visam limitar a multicolinearidade em um modelo preditivo da densidade hepática a partir de características radiômicas extraídas de exames de TC.

Metodologia

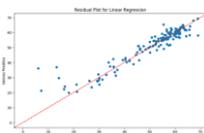
Foi conduzida uma pesquisa retrospectiva observacional com base em TC abdominais de 450 pacientes. As imagens foram processadas para a extração de características radiômicas, utilizando-se o Software Slicer, para segmentação hepática e posterior coleta de dados quantitativos. Foram testados modelos preditivos com o objetivo de redução de multicolinearidade.

Resultados

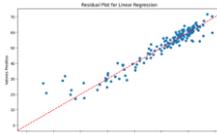
MSE = 21,6, indicando boa acurácia, porém com multicolinearidade severa ($VIF > 100$ mil), comprometendo a interpretação dos coeficientes.



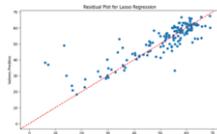
LASSO manteve boa acurácia ($R^2 = 0,88$; MSE = 22,1) e reduziu parcialmente a multicolinearidade, embora $VIFs > 200$ mil ainda limitem a interpretação dos coeficientes.



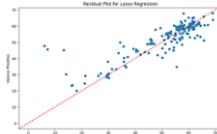
RIDGE reduziu parcialmente a multicolinearidade ($VIFs > 4$ mil) e apresentou melhor desempenho preditivo ($R^2 = 0,91$; MSE = 16,6), superando os modelos anteriores em precisão.



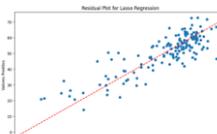
A seleção por $VIF < 10$ eliminou a multicolinearidade, mas reduziu a acurácia preditiva ($R^2 = 0,73$; MSE = 47,7), indicando perda de desempenho ao priorizar estabilidade.



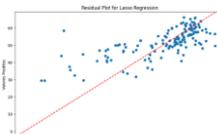
A combinação de LASSO + VIF eliminou a multicolinearidade ($VIF < 10$), mas reduziu a precisão preditiva ($R^2 = 0,70$; MSE = 53,9), comprometendo o desempenho do modelo.



RIDGE + VIF reduziu a multicolinearidade ($VIF < 10$) e teve desempenho moderado ($R^2 = 0,71$; MSE = 51,6), equilibrando estabilidade estatística e precisão preditiva.



PCA eliminou totalmente a multicolinearidade por meio de componentes ortogonais, mas teve a menor precisão preditiva ($R^2 = 0,50$; MSE = 89,7) entre os modelos avaliados.



Conclusão

Este estudo evidenciou que a multicolinearidade nas características radiômicas afeta a precisão dos modelos preditivos de gordura hepática. Estratégias como LASSO e RIDGE reduziram esse efeito, mas com limitações na estabilidade dos coeficientes. A combinação RIDGE + VIF apresentou melhor equilíbrio entre precisão e estabilidade. PCA eliminou multicolinearidade, porém comprometeu a interpretabilidade. A escolha adequada da técnica é essencial para diagnósticos confiáveis e não invasivos.