

Núcleo de Avaliação: Núcleo II

Área temática: Ciências Exatas e da Terra

Área do Conhecimento: Ciência da Computação

Estudo e aplicação de técnicas de inteligência computacional para análise da saúde da mulher

Wesley dos Santos Silva, Samara Martins Nascimento Gonçalves

A Síndrome dos Ovários Policísticos (SOP) é um distúrbio endócrino que afeta mulheres em idade reprodutiva e de difícil diagnóstico devido à sua heterogeneidade clínica e sobreposição de sintomas com outras condições, como resistência à insulina e obesidade (Rangel et al., Brazilian Journal of Implantology and Health Sciences, 6:5403–541, 2024). Em virtude disso, estratégias computacionais surgem como uma abordagem promissora para aprimorar a precisão do diagnóstico dessa síndrome (Suha; Islam, Heliyon, 9:e14518, 2023). Nesse sentido, o objetivo deste estudo consiste na aplicação e utilização de Machine Learning (ML) com a finalidade de desenvolver uma ferramenta que auxilie no diagnóstico da SOP. Para isso, foi utilizado um conjunto de dados clínicos de 541 mulheres, com faixa etária entre 20 a 48 anos, de diferentes hospitais da Índia, obtidos na plataforma digital Kaggle. Esses dados passaram por um pré-processamento, que incluiu desde o tratamento de dados ausentes até a aplicação de técnicas de seleção de características e balanceamento de dados. Os modelos empregados neste estudo foram os seguintes: Support Vector Machine (SVM), modelo que busca a maximização de margens; K-Nearest Neighbors (KNN), um modelo que se baseia na distância entre os objetos para classificar novos dados; e, por fim, Random Forest (RF) e Decision Tree (DT), que são modelos simbólicos e respaldados na estratégia de dividir para conquistar, sendo que o primeiro é formado por um conjunto do segundo. Tais modelos foram avaliados mediante as métricas Acurácia, Precision, Recall e F1-Score, que medem a proporção, qualidade e a capacidade dos modelos em lidar com casos positivos e negativos no conjunto de dados. Após o treinamento e avaliação dos modelos, o RF apresentou o melhor desempenho em termos de Acurácia e F1-Score, com 93% em ambas as métricas, sinalizando a sua robustez na identificação de casos positivos e negativos de SOP. Com base nesse resultado, foi desenvolvida a SOP ASSIST, uma Application Programming Interface (API) implementada utilizando o framework Flask e que disponibiliza o diagnóstico preditivo da paciente utilizando o modelo otimizado (RF). Após testes funcionais e preditivos considerando amostras sintéticas positivas e negativas, foi possível verificar que API processa corretamente as requisições, fornecendo respostas rápidas e coerentes com as expectativas. Dessa forma, apesar dos desafios na identificação da SOP, o uso de ML pode contribuir no diagnóstico da doença. O modelo mais robusto nas métricas de avaliação (RF) aliado à SOP ASSIST (API), formam uma ferramenta de apoio à decisão clínica, com impacto social positivo para a saúde feminina.

Palavras-chave: Síndrome dos Ovários Policísticos, Saúde Reprodutiva, Machine Learning.

ncia financiadora: PICI-UFERSA

Campus: Angicos

