

ALGORITMO DE VALIDAÇÃO PARA PREDIÇÃO DE FALSOS POSITIVOS NA TRIAGEM NEONATAL DE FIBROSE CÍSTICA

Custódio PRS¹, Almeida RMS¹, Oliveira VK²

¹ Intercientífica, Assistência Técnica, Parque Tecnológico UNIVAP, paulo55866@gmail.com.br.

² Instituto de Pesquisa e Desenvolvimento (IP&D) UNIVAP, Laboratório de Física e Astronomia.

Resumo

Introdução. O diagnóstico sintomático de fibrose cística (FC) traz complicações pulmonares, metabólicas, gastrointestinais, isso por conta do aumento da produção de muco com características mais espessas do que a da população que não tem a doença. Tendo como principal tratamento indicado o pré sintomático para tardar os efeitos da doença. Esta doença atinge com maior frequência a população branca sem diferença entre sexos e um estudo realizado na França em 2002 constatou que a população de origem africana sofre de incidência expressivamente menor do que na população caucasiana, sendo assim na maioria dos casos triados positivamente são falsos positivos (FP). Além desses aspectos sabe-se que falsos positivos podem se dar por conta de erros laboratoriais, mudança de procedimento, ou fatores biológicos. Estudos mostram que resultados FP podem causar impactos psicológicos na família, como ansiedade e angústia e isso pode ser amenizado da forma que a abordagem aos pais é feita. **Objetivos.** Desenvolver um algoritmo de Aprendizado de Máquina (AM) para prever FP na triagem de fibrose cística, utilizando uma mescla de dados reais e dados simulados. O modelo deve ser capaz de criar critérios e encontrar relações entre essas características e o diagnóstico. **Metodologia.** Foram utilizados dados gerais sobre quantidade de pacientes triados positivos, totais e confirmado para FC, fornecidos por um dos clientes da Intercientífica, juntamente de características dos pacientes que tiveram valores simulados, para desenvolver um modelo de florestas aleatórias. **Resultados.** O primeiro modelo de validação se mostrou eficiente tendo acurácia e sensibilidade de 100% e precisão de 75% nas previsões realizadas em um conjunto de dados desconhecidos pelo modelo de treino. **Conclusão.** Por se tratar de um modelo de validação da teoria de um projeto que utilizará dados reais para fazer previsão essa sistemática adotada mostra que pode ser possível fazer previsões levando em conta as características do paciente e com isso auxiliar na abordagem ao país após a triagem.

Palavras-chave: Fibrose Cística, Florestas aleatórias, triagem neonatal.

Áreas de Concentração: Desenvolvimento de Métodos e Dispositivos Diagnósticos