

USO DE OTIMIZAÇÃO E VISÃO COMPUTACIONAL PARA CLASSIFICAÇÃO DE DENGUE COM DADOS TABULARES DO SUS

Danielle Vasconcelos Gomes

FCA - UNICAMP

Limeira - São Paulo

d245465@dac.unicamp.br

Cristiano Torezzan

FCA - UNICAMP

Limeira - São Paulo

torezzan@unicamp.br

RESUMO

Este trabalho apresenta uma abordagem híbrida que combina técnicas de otimização e visão computacional para a classificação automática de casos de dengue a partir de dados tabulares oriundos de registros públicos do Sistema Único de Saúde (SUS). Embora redes neurais convolucionais (CNNs) sejam altamente eficazes em tarefas de classificação de imagens, sua aplicação direta a dados tabulares — comuns em contextos de saúde pública — é desafiadora, dada a ausência de estrutura espacial explícita nesses dados. Para contornar essa limitação e explorar o potencial das CNNs, trabalhos recentes têm proposto a transformação de vetores de atributos em imagens bidimensionais, preservando relações locais entre as variáveis. A transformação utilizada neste trabalho é baseada na técnica t-Distributed Stochastic Neighbor Embedding (t-SNE), que utiliza um processo iterativo de otimização da divergência de Kullback–Leibler para projetar os dados em um espaço visual coerente com sua estrutura original. As imagens geradas são então utilizadas como entrada para CNNs treinadas para classificar notificações como casos positivos ou negativos de dengue. Os experimentos realizados com dados reais do Open DataSUS mostram que a abordagem proposta pode superar modelos clássicos de aprendizado supervisionado na tarefa de predição do diagnóstico. Esses resultados reforçam o potencial do uso combinado de técnicas de otimização e aprendizado profundo na construção de ferramentas para apoio à vigilância epidemiológica e à tomada de decisão em saúde pública.

PALAVRAS CHAVE. Dados Tabulares; Otimização; Visão Computacional; Dengue.

ABSTRACT

This work presents a hybrid approach that combines optimization techniques and computer vision for the automatic classification of dengue cases based on tabular data from public records of the Brazilian Unified Health System (SUS). Although convolutional neural networks (CNNs) are highly effective in image classification tasks, their direct application to tabular data—common in public health contexts—is challenging due to the lack of explicit spatial structure in such data. To overcome this limitation and leverage the potential of CNNs, recent studies have proposed transforming feature vectors into two-dimensional images while preserving local relationships among variables. The transformation used in this study is based on the t-Distributed Stochastic Neighbor Embedding (t-SNE) technique, which employs an iterative optimization process to minimize the Kullback–Leibler divergence between the data distributions in the original and projected spaces.

The resulting images are then used as input to CNNs trained to classify notifications as positive or negative dengue cases. Experiments conducted with real data from Open DataSUS show that the proposed approach can outperform classical supervised learning models in diagnostic prediction tasks. These results reinforce the potential of combining optimization techniques and deep learning in the development of tools to support epidemiological surveillance and public health decision-making.

KEYWORDS. Tabular Data; Optimization; Computer Vision; Dengue.