

## Utilização de *machine learning* para classificação de plantas saudáveis e plantas submetidas ao estresse

João Victor Silva Viana<sup>1</sup>; Leticia Piazzentin Dantas<sup>2</sup>; Otávio Rodrigues de Paula<sup>3</sup>; Bianca Batista Barreto<sup>4</sup>; Débora Marcondes Bastos Pereira<sup>5</sup>

<sup>1</sup>Aluno de graduação em Física Computacional, Universidade de São Paulo, São Carlos, SP. Bolsista CNPq, Embrapa Instrumentação, São Carlos, SP; vianajoao@usp.br

<sup>2</sup>Aluna de graduação em Biotecnologia, Universidade Federal de São Carlos, São Carlos, SP.

<sup>3</sup>Aluno de graduação em Biotecnologia, Universidade Federal de São Carlos, São Carlos, SP.

<sup>4</sup>Pós doutoranda da Embrapa Instrumentação, São Carlos, SP.

<sup>5</sup>Pesquisadora da Embrapa Instrumentação, São Carlos, SP (orientadora).

A integração de tecnologias na agricultura tem se tornado cada vez mais comum, trazendo consigo uma série de benefícios, desde o aumento da produtividade até contribuições para a sustentabilidade. Uma das tecnologias promissoras que pode ter um papel importante nesse cenário é o aprendizado de máquina (*machine learning*), especialmente no que diz respeito à avaliação da saúde das plantas. Neste contexto, o objetivo do trabalho foi classificar plantas saudáveis e inoculadas com patógeno utilizando quatro diferentes técnicas de machine learning: Análise Discriminante, Regressão Logística, *AdaBoost*, um algoritmo de *boosting*, que combina múltiplos classificadores fracos para formar um classificador mais preciso, e SVM (*Support Vector Machine*), que é um algoritmo que separa os dados usando um hiperplano. Para esse propósito, foram utilizados dados coletados de três técnicas distintas: Fluorcam, Espectroscopia induzida por laser (LIFS) e Câmera Térmica. No âmbito computacional, diferentes modelos foram treinados e testados para cada equipamento. Os resultados revelaram que a melhor taxa de acerto foi obtida com os dados de imagem de fluorescência fornecidos pelo equipamento Fluorcam, alcançando uma taxa de acerto de 75% utilizando o modelo SVM. Para o LIFS, a taxa de acerto foi maior, atingindo 80% também utilizando o modelo SVM. Já para os dados de termografia a taxa de acerto foi de 90% utilizando o *AdaBoost*. Esses resultados evidenciam a relevância e o potencial do uso do machine learning na classificação de plantas inoculadas, especialmente quando combinado com tecnologias específicas. A partir desses resultados, conclui-se que o desenvolvimento de soluções usando *machine learning* pode contribuir significativamente para o monitoramento e aprimoramento da detecção do ataque de patógenos na planta, impulsionando assim a eficiência e a sustentabilidade da agricultura.

**Apoio financeiro:** CNPq

**Área:** Ciências Exatas e da Terra

**Palavras-chave:** *machine learning*, classificação, detecção de doenças

**N. do Processo PIBIC/PIBITI:** 372370/2022-5