

## Resumo

A produção de cana-de-açúcar no Brasil registrou um faturamento de US\$18,27 bilhões no ano comercial da safra 2023/2024, com destaque para o estado de São Paulo, líder na produção sucroalcooleira do país. Embora seja uma cultura amplamente difundida na economia, prever a produção da cana-de-açúcar é algo extremamente complexo, principalmente devido aos diversos fatores que influenciam esse resultado, como condições climáticas variáveis, práticas agronômicas específicas e características do solo. Diante desse cenário, este estudo propõe uma análise de algoritmos de machine learning com o objetivo de avaliar os seus desempenhos na previsão da produção de cana-de-açúcar em São Paulo. Para isso, foi criada uma base de dados robusta, composta por informações históricas sobre safras anteriores e variáveis influenciadoras, como clima e solo. Os modelos foram avaliados através de métricas de desempenho, como acurácia, precisão, recall, F1-score, erro médio absoluto e raiz quadrada do erro quadrático médio. Nossos resultados sugerem que a Floresta Aleatória foi o modelo mais eficaz para prever a produção de cana-de-açúcar, embora o Support Vector Machine também tenha obtido um bom desempenho. Este trabalho contribui para a aplicação de métodos de previsão no setor agrícola, oferecendo uma abordagem sistemática para aprimorar a previsão de produção de cana-de-açúcar. Assim, a pesquisa não apenas reforça a importância da análise de dados para uma agricultura mais precisa e eficiente, mas também serve como referência para futuras iniciativas que busquem melhorar a gestão e a sustentabilidade do setor sucroalcooleiro.



