**Optimizing Crop Yield Prediction: A Comparative Analysis of Machine Learning Methods**

**Authors:** Dina Dinh**,** Thanos Gentimis, Yinan Yang, Leticia Santos

**Abstract**

**Objective:** The main purpose of this study is to compare machine learning methods and observe the disparities between optimized and unoptimized models for yield prediction in Soybeans.

**Method:** The methodology involves the utilization and optimization of neural networks using a Python implementation of Tensorflow with Scikit-learn and Keras. We propose a dense neural network construction, determining the optimal number of neurons in each hidden layer and the necessary number of hidden layers, guided by a specific statistically motivated criterion. Additionally, we propose the implementation and optimization of a random forest as a standard comparison model.

**Conclusion:** The findings of this study highlight the importance of optimization in enhancing the efficacy of machine learning models for predicting crop yield. The neural network is the preferred model for this dataset since the dataset is large and complex. Given the dataset's complexity, no discernible pattern can be identified, further emphasizing the need for advanced modeling techniques, such as the optimized neural network that can adapt well to this dataset.