Development of a Hybrid Machine Learning Tool for Rainfall Prediction Based on the Traditional Knowledge of Rain Prophets and Meteorological Data

Magalhaes Souza, Raul (School: Escola de Ensino Medio Joaquim de Figueiredo Correia)

Long-term rainfall prediction in the Brazilian semi-arid region remains a significant challenge for conventional meteorological systems. These systems are often expensive and inaccurate at local scales due to the high variability associated with the Intertropical Convergence Zone (ITCZ), creating critical uncertainties for agriculture-dependent communities. Addressing this scenario, this research presents a hybrid machine learning model integrating 44 years of meteorological data (1981–2024; >40,000 records) from the Jaguaribe Valley, Brazil, with the ancestral knowledge of "rain prophets," structured into botanical, faunal, and atmospheric indicators. Several algorithms were developed and trained to effectively capture complex non-linear patterns. During validation in the 2025 rainy season (January to March), the best model achieved exceptional performance: it predicted 422.1 mm of accumulated rainfall against 431.5 mm observed, resulting in a mean monthly error of only 5.7%. In contrast, a leading commercial platform underestimated the volume (318 mm), presenting a mean error of 30.8%. Regarding classification performance, the system achieved 94.5% accuracy, an F1-score of 94%, and an AUC-ROC of 96%, consistently outperforming purely meteorological methods. Beyond high precision, the model is computationally lightweight and low-cost, enabling its deployment in rural communities. It is concluded that integrating data science with traditional knowledge consolidates an innovative, scalable, and robust social technology for resilient agriculture and water management in drought-affected regions.

Awards Won: