ECP 2024 Conference
The 3rd International Electronic Conference on Processes
29–31 May 2024 | Online

Harvesting Insights: AI-driven Rice Yield Predictions and Big Data Analytics in Agriculture

Cheikh Abdelkader Ahmed Telmoud
Scientific Computing, Computer Science and Data Science Research Unit (CSIDS), University of Nouakchott, Nouakchott, Mauritania

INTRODUCTION & AIM

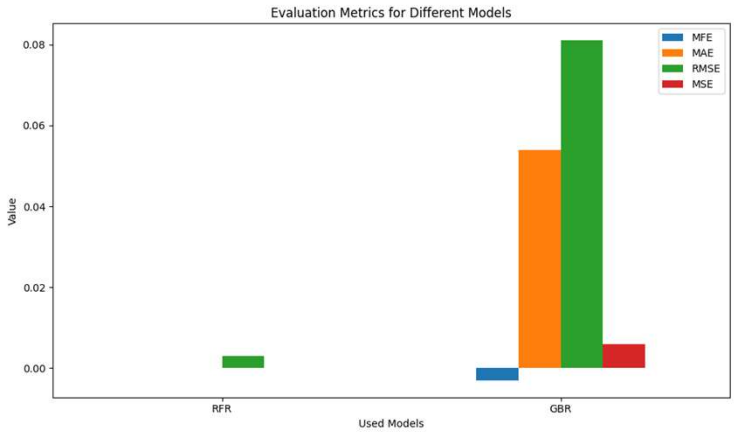
This paper explores the transformative potential of artificial intelligence (AI) and big data analytics in predicting rice yields within the agricultural domain. By employing advanced algorithms and innovative methodologies, our study aims to contribute to the optimization of crop management strategies, providing a glimpse into the future of sustainable agriculture. The integration of AI and big data analytics allows us to unveil novel insights into rice yield predictions, emphasizing their broader implications for global food security.

METHOD

Our innovative methodologies consist of combining climate data and cultivated areas as input to predict rice yield for a season based on these critical factors. We employ well-known models from the literature, such as Random Forest Regression (RFR) and Gradient Boosting Regression (GBR), augmented with advanced optimization methods such as Optuna and Hyperopt, to enhance our model. The integration of AI with big data analytics in rice yield predictions aids in preparing the data to achieve high quality before applying our models. This includes selecting optimal features and simulating our model with generated data to ensure it avoids overfitting.

We have used a national dataset collected from Rosso, Mauritania, a region of significant agricultural importance in Africa. A critical water source for agriculture in Mauritania is the Senegal River, serving as the primary drinking water source for over 4.65 million individuals (according to 2020 data from the Ministry of Housing and Urban Development of Mauritania).

RESULTS & DISCUSSION



Our optimized Random Forest Regression model exhibited impressive results, with a Mean Forecasting Error (MFE) of 0.0001, a Mean Absolute Error (MAE) of 0.00016, a Mean Square Error (MSE) of 0.000014, and a Root Mean Square Error (RMSE) of 0.003. These results strongly endorse the efficacy of our decision to employ Random Forest Regression (RFR) for yield prediction, demonstrating its superior performance compared to other models used in the literature.

CONCLUSION

Finally, the potential of our approach has led to the creation of a digital agricultural twin for monitoring, analyzing, and visualizing data provided by sensors installed on a farm in Rosso during the study period.

FUTURE WORK / REFERENCES

As a future perspective, we aim to create an intelligent system using our models and integrate IoT technologies to expose our model results as a service. We plan to publish the first Mauritanian agricultural database for other researchers to use in their future research.