Diverse Strategies for Enhancing Agricultural Yield in India

1st Ana Moreira
MEIA
ISEP
Aveiro, Portugal
1222731@isep.ipp.pt

2nd Carlos Dias
MEIA
ISEP
Porto, Portugal
1200991@isep.ipp.pt

3rd Diana Magalhães
MEIA
ISEP
Porto, Portugal
1201830@isep.ipp.pt

4th Diogo Violante
MEIA
ISEP
Porto, Portugal
1201284@isep.ipp.pt

5th Gustavo Ferreira
MEIA
ISEP
Porto, Portugal
1200829@isep.ipp.pt

***Abstract*—This paper explores into the integration of machine learning (ML) in agriculture, emphasizing productivity and sustainability. It provides a concise overview of existing research, spotlighting ML algorithms in precision agriculture. The section on planning and decision-making explores the potential of different algorithms in this work. This paper contributes to the growing knowledge on ML applications in agriculture, presenting diverse methodologies for precision farming and yield prediction.**

***Index Terms*—Agriculture, Machine learning, Data Analysis, Prediction, India, Crop**

I. Introduction

In the realm of agriculture, the integration of machine learning has become a pivotal area of study. This introduction surveys existing researches, focusing on different algorithms and its applications in farming and crop optimization. Our contribution to this field is also outlined, emphasizing the seamless incorporation of machine learning into agricultural practices.

II. State of the Art

A. Agriculture

Agriculture, a cornerstone of our global economy, is undergoing a transformative evolution through precision agriculture, also known as digital farming or intelligent agriculture, this approach integrates some technologies into traditional farming methods [1]. The objective is clear: enhance agricultural productivity and sustainability to meet the rising demands of a growing human population.

In response to these challenges, agri-technology and precision farming have emerged as dynamic scientific fields. Termed as digital agriculture, they leverage data-intensive strategies to optimize productivity while minimizing environmental impact [2]. The machinery data contributes to more informed and swift decision-making processes.

Agriculture, a sector traditionally rooted in manual labor, is experiencing a transformative revolution through the integration of Machine Learning [3].

B. Machine Learning

As a subset of Artificial Intelligence, Machine Learning empowers computers to emulate human thinking patterns, evolving from different experiences [3]. This technological leap has been paralleled by advancements in storage and processing power, ushering in innovative applications in diverse fields, from self-driving cars to personalized recommendations on different platforms [4].

In the realm of agriculture, Machine Learning emerges as a powerful tool for optimizing processes and enhancing decision-making. By harnessing statistical methods and algorithms, Machine Learning facilitates classifications, predictions, and insightful data mining within the agricultural landscape [4]. The potential applications range from crop yield predictions and resource optimization to disease detection in crops. As our world becomes increasingly data-driven, the integration of Machine Learning in agriculture not only improves efficiency but also holds the promise of addressing key challenges in food production [3].

C. Related Work

This section aims to analyze three distinct works related to our research in the intersection of machine learning and agriculture. Following the individual analyses, a comprehensive comparison will be provided, offering insights into the methodologies employed in each of these works.

1) Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: The principal aim of this review paper is to assess and discuss the various ML approaches employed in precision agriculture for crop yield prediction and nitrogen status estimation. The study critically examines advancements in remote sensing systems and underscores the necessity for ML methods to handle the substantial volume of data generated. The paper emphasizes several ML techniques, including Back-propagation Neural Networks, Convolutional Neural Networks, probabilistic Processes, M5-Prime Regression Trees, Least Squares Support Vector Machines, and Fuzzy Cognitive Maps, showcasing their successful applications across diverse precision agriculture tasks [5].

Back-propagation Neural Networks (BNN) play a pivotal role in identifying the importance of various Vegetation Indices (VI), enhancing the precision of crop yield estimation. Combining Convolutional Neural Networks (CNN) and Long-short Term Memory (LSTM) with Gaussian Processes proves effective in extracting features from data and minimizing error maps. Gaussian Processes (GP) contribute by automatically selecting wavebands across the spectrum, facilitating the prediction of different characteristics of plant leaves. M5-Prime Regression Trees emerge as a suitable tool for multi-class crop prediction, showcasing versatility in agricultural applications. Least Squares Support Vector Machine (LS-SVM) stands out as a promising ML technique for regression analysis, specifically in quantifying nitrogen status. Additionally, Fuzzy Cognitive Map (FCM) finds application in modeling and representing expert knowledge, particularly in the context of yield prediction and crop management. Together, these ML techniques offer valuable tools for advancing precision agriculture practices.

In summary, the reviewed paper provides valuable insights into the role of machine learning in enhancing crop yield prediction and nitrogen management in precision agriculture. It sheds light on the advantages of employing ML techniques in remote sensing for agriculture and anticipates future trends involving sensor optimization, expert knowledge integration, and the synergy of diverse ML techniques.

2) Machine learning techniques for forecasting agricultural prices: In this section, it will be analyzed a work wit the title ”Machine learning techniques for forecasting agricultural prices: A case of brinjal in Odisha, India”, that makes use of several machine learning algorithms, to forecast the wholesale price of brinjal in the seventeen major markets of Odisha in India [6].

Algorithms such as Generalized Neural Network (GRNN), Support Vector Regression (SVR), Random Forest (RF) and Gradient Boosting Machine (GBM) were used in this paper, given that they are considered to be efficient regarding the scope of the study, and were compared with the performance of a Autoregressive Integrated Moving Average (ARIMA) model. This was achieved with the use of Model Confidence Set (MCS), Mean Error (ME), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Prediction Error (MAPE).

The dataset was divided in two, 90% used for model development and the other 10% for validation purposes and the comparison between each model’s prediction was achieved by performing a Diebold-Mariano test, to observe the differences between algorithms and to further cement the superiority of certain algorithms circular and radar plots were also used.

It is concluded that the ML algorithms used had a better performance than the ARIMA, with GRNN having the better performance between them, setting the relevance of the use of these models in future problems related to the topic explored. Still, with this proven, there are several factors that were not taking into consideration, which could affect the prediction accuracy of the models and those should be taken into account.

3) Crop Yield prediction in different models: The work analyzed in this section is titled ”A Comparison Between Major Artificial Intelligence Models for Crop Yield Prediction: Case Study of the Midwestern United States, 2006–2015” [7].

The primary objective of this paper is to evaluate and compare various Artificial Intelligence (AI) models to determine the most effective one for predicting crop yield in the Midwestern United States. The study delves into the analysis of a comprehensive data set, encompassing the characteristics of the study area, meteorological and hydro-logical data, cropland data, and cropland yield statistics spanning the period from 2006 to 2015.

Six distinct models were developed using the created data set, namely: Multivariate Adaptive Regression Splines (MARS), Support Vector Machine (SVM), Random Forest (RF), Extremely Randomized Trees (ERT), Artificial Neural Network (ANN), and Deep Neural Network (DNN). Given the annual nature of crop harvesting, these models were employed to predict and validate yields on a yearly basis.

Upon conducting an in-depth analysis of the correlation matrix derived from the data set, the models were constructed, revealing that the DNN model exhibited the lowest Mean Average Error (MAE). The MAE for the DNN model proved to be 21% to 33% more accurate than the other models considered.

Conclusively, the research paper asserts that the DNN model can be reasonably relied upon to predict future yield outcomes with a degree of reliability. This finding suggests the practical applicability of the DNN model in forecasting crop yields in subsequent years.

4) Related Work Comparison: All three works share the common objective of applying machine learning techniques in agriculture, addressing specific challenges such as crop yield prediction, nitrogen status estimation, and agricultural price forecasting. Each work uses the different machine learning techniques for its specific context, showing the versatility of machine learning in addressing different agricultural challenges.

There is a recurring theme in the works with the use of Neural Networks, with BNN, GRNN, and DNN being highlighted as effective models. Each work underscores the superiority of ML techniques over traditional methods, showcasing the potential for increased accuracy and reliability in agricultural predictions.

In summary, these works collectively contribute to the growing body of knowledge on the application of ML in agriculture, highlighting diverse methodologies and models that can be employed to enhance precision farming, price forecasting, and yield prediction.

D. Planning and Decision Making

Collaborating seamlessly with machine learning [8] or operating independently, decision-making algorithms have the potential to enhance the efficacy of a given model or streamline the output of a specific system [9].

As the name suggests, decision-making algorithms are specifically crafted to assist in selecting the most appropriate course of action in situations characterized by uncertainty. This includes applications ranging from route planning to disaster response strategies.

When focusing on the agricultural domain, planning and decision-making algorithms emerge as formidable tools with the capability to optimize not only crop production but also associated costs [10]. Their utilization extends to tasks such as path planning for field machines [11] and crop selection [12].

E. Related Work

This section is dedicated to examining three specific works pertinent to our research at the crossroads of decision making and planning. Subsequent to the individual analyses, a comprehensive comparison will be presented, delving into the methodologies employed in each of these works, with a specific focus on understanding the algorithms utilized in each one.

1) Genetic Algorithm Applied to Weather Forecast and Agricultural Decision Making Optimization: A genetic algorithm is based on biological evolution, it simulates the natural selection of a certain population of individuals, which allows to generate optimized solutions of the initially posed problem. This is achieved by selecting individuals from an initial population that are more fit and this process repeats, while also taking into consideration the probability of mutations and crossovers [13].

The application of Genetic Algorithms in Agriculture can have great benefits, for example in decision making, has we can see in this paper by S.K. Roy and D. De, where this type of algorithm is used to optimize decision making in agriculture, by improving the predictions of rainfall. They achieve this by using a real weather data set and when the system fails to make a prediction, a sensor-based watering system is activated. In this paper, two lemmas and a theorem are defined, where the first two explicate probabilities related to the population and the theorem is related to the fitness of the population in a given interaction. They conclude that the system is beneficial for farmers that have drought-affected areas, to make the most of their water distribution [14].

As so, Genetic Algorithms can prove beneficial for projects that have the need for high-quality solutions for optimization and search problems [15].

2) Simulated Annealing and On-Farm Irrigation Optimization: Simulated Annealing, a global search optimization algorithm, gradually refines solutions by mimicking metallurgical annealing. Its probabilistic acceptance of sub-optimal solutions suits the uncertainty in agricultural variables [16]. The algorithm's adaptability, driven by the Boltzmann probability distribution, makes it well-suited for escaping local minima [17].

In the realm of on-farm irrigation scheduling, the Canterbury irrigation scheduler(CIS) leverages Simulated Annealing to maximize farm profit in water-limited scenarios [18]. By adjusting irrigation decisions based on detailed farm system models, including realistic plant responses, the CIS outperforms traditional stochastic dynamic programming schedulers [18].

The CIS method, optimized with Simulated Annealing, excels in providing near-optimal decisions and showcasing adaptability to realistic farm models. In a case study, it demonstrated a potential 10% increase in pasture yield revenue in Canterbury compared to conventional scheduling practices [18].

3) Ant Colony Optimization for Agricultural Contracting Work Scheduling Optimization: In this section, we delve into a research paper that focuses on optimizing the scheduling of agricultural contracting work in Finland with an algorithm based in the Ant Colony Optimization (ACO) Algorithm with modifications.

ACO draws inspiration from the behavior of real ants when searching for food. Initially ants explore around their nest in a seemingly random manner. Upon discovering a food source, an ant assesses both the quantity and quality of the available sustenance and then transports a portion of the food back to the nest. During the return journey the ant leaves behind a pheromone trail on the ground. This deployment of pheromones serves as a form of communication among the ant colony. The indirect communication mechanism via pheromone trails enables the ants to collaboratively discover and reinforce the shortest paths between their nest and various food sources [19].

In the specific context of agricultural contracting, the work presented proposes an ACO-based algorithm for scheduling tasks. The model considers factors such as task locations, time windows, and resource requirements, reflecting the real-world constraints faced by contractors. The algorithm is tailored to the unique characteristics of agricultural contracting, where tasks are known in advance, and the optimization of travel routes is essential to minimize costs [20].

The paper presents the results of benchmarking the algorithm on known instances of the Traveling Salesman Problem (TSP) and randomized instances. Performance metrics, including path length and solution lateness, demonstrate the algorithm's effectiveness in producing high-quality solutions for instances of practical size [20].

F. Algorithms Comparison

Genetic Algorithms, Simulated Annealing, and Ant Colony Optimization are all meta heuristic algorithms used in agricultural optimization. Genetic Algorithms are based on natural selection principles and are good at generating optimized solutions through simulated evolution. In the other hand, Simulated Annealing is inspired by metallurgical annealing and is good at escaping local minima. Lastly, Ant Colony Optimization is based on pheromone communication and is good at collaborative exploration.

Each algorithm has its own strengths and weaknesses, and the best algorithm to use for a particular problem will depend on the specific details of the problem. Genetic Algorithms are well-suited to problems with a large number of possible solutions, while Simulated Annealing is good for problems with a large number of local minima. Ant Colony Optimization is good for problems where there is a lot of uncertainty.

References

[1] S. Condran, M. Bewong, M. Z. Islam, L. Maphosa, and L. Zheng, "Machine learning in precision agriculture: a survey on trends, applications

and evaluations over two decades," *IEEE Access*, vol. 10, pp. 73 786–73 803, 2022.

[2] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, p. 2674, 2018.

[3] S. V. Mahadevkar, B. Khemani, S. Patil, K. Kotecha, D. R. Vora, A. Abraham, and L. A. Gabralla, "A review on machine learning styles in computer vision—techniques and future directions," *IEEE Access*, vol. 10, pp. 107 293–107 329, 2022.

[4] IBM, "What is machine learning?" last accessed 08 Dec 2023. [Online]. Available: https://www.ibm.com/topics/machine-learning

[5] A. Chlingaryan, S. Sukkarieh, and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review," *Computers and Electronics in Agriculture*, vol. 151, pp. 61–69, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168169917314710

[6] R. K. Paul, M. Yeasin, P. Kumar, P. Kumar, M. Balasubramanian, H. Roy, A. Paul, and A. Gupta, "Machine learning techniques for forecasting agricultural prices: A case of brinjal in odisha, india," *Plos one*, vol. 17, no. 7, p. e0270553, 2022.

[7] N. Kim, K.-J. Ha, N.-W. Park, J. Cho, S. Hong, and Y.-W. Lee, "A comparison between major artificial intelligence models for crop yield prediction: Case study of the midwestern united states, 2006–2015," *ISPRS International Journal of Geo-Information*, vol. 8, no. 5, p. 240, 2019.

[8] M. Elbasheer, F. Longo, L. Nicoletti, A. Padovano, V. Solina, and M. Vetrano, "Applications of ml/ai for decision-intensive tasks in production planning and control," *Procedia Computer Science*, vol. 200, pp. 1903–1912, 2022, 3rd International Conference on Industry 4.0 and Smart Manufacturing. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050922004008

[9] M. J. Kochenderfer, T. A. Wheeler, and K. H. Wray, *Algorithms for decision making*. MIT press, 2022.

[10] K. Alibabaei, P. D. Gaspar, T. M. Lima, R. M. Campos, I. Girão, J. Monteiro, and C. M. Lopes, "A review of the challenges of using deep learning algorithms to support decision-making in agricultural activities," *Remote Sensing*, vol. 14, no. 3, p. 638, 2022.

[11] T. Oksanen and A. Visala, "Coverage path planning algorithms for agricultural field machines," *Journal of field robotics*, vol. 26, no. 8, pp. 651–668, 2009.

[12] N. Deepa and K. Ganesan, "Decision-making tool for crop selection for agriculture development," *Neural Computing and Applications*, vol. 31, pp. 1215–1225, 2019.

[13] S. Forrest, "Genetic algorithms," *ACM computing surveys (CSUR)*, vol. 28, no. 1, pp. 77–80, 1996.

[14] S. K. Roy and D. De, "Genetic algorithm based internet of precision agricultural things (iopat) for agriculture 4.0," *Internet of Things*, vol. 18, p. 100201, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660519300071

[15] K. Deb, "Genetic algorithm in search and optimization: the technique and applications," in *Proceedings of International Workshop on Soft Computing and Intelligent Systems,(ISI, Calcutta, India)*. Proceedings of International Workshop on Soft Computing and Intelligent . . . , 1998, pp. 58–87.

[16] J. Brownlee, "Simulated annealing from scratch in python," 2021, last accessed 10 Dec 2023. [Online]. Available: https://machinelearningmastery.com/simulated-annealing-from-scratch-in-python/

[17] Baeldung, "Simulated annealing explained," 2023, last accessed 10 Dec 2023. [Online]. Available: https://www.baeldung.com/cs/simulated-annealing

[18] P. D. Brown, T. A. Cochrane, and T. D. Krom, "Optimal on-farm irrigation scheduling with a seasonal water limit using simulated annealing," *Agricultural Water Management*, vol. 97, no. 6, pp. 892–900, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378377410000387

[19] M. Dorigo and C. Blum, "Ant colony optimization theory: A survey," *Theoretical computer science*, vol. 344, no. 2-3, pp. 243–278, 2005.

[20] S. Alaiso, J. Backman, and A. Visala, "Ant colony optimization for scheduling of agricultural contracting work," *IFAC Proceedings Volumes*, vol. 46, no. 18, pp. 133–137, 2013.