Maximizing Agricultural Productivity: A Comprehensive Survey and Analysis of Crop Recommendation Systems

Mohd Rafi Lone
VIT Bhopal University, Bhopal, India
mohdrafilone@vitbhopal.ac.in

KV Monish
VIT Bhopal University, Bhopal, India
kv.monish2020@vitbhopal.ac.in

Aryan Kashyap
VIT Bhopal University, Bhopal, India
aryan.kashyap2020@vitbhopal.ac.in

Aryan Singh Rajput
VIT Bhopal University, Bhopal, India
aryan.singh2020@vitbhopal.ac.in

Shrey Khanduja
VIT Bhopal University, Bhopal, India
shrey.khanduja2020@vitbhopal.ac.in

Sagar Maheshwari
VIT Bhopal University, Bhopal, India
sagar.maheshwari2020@vitbhopal.ac.in

***Abstract*—Effective decision-making in agriculture, particularly in crop selection, is pivotal for maximizing yields and economic gains. This survey investigates the application of diverse machine learning algorithms, including Artificial Neural Networks (ANN), in customizing crop recommendation systems for soil features. Focusing on identifying the optimal approach, the survey critically examines various recommender systems, highlighting strengths and drawbacks in alternative algorithms and methodologies. Through a comparative analysis, it seeks to guide farmers in making informed decisions about crop selection, ultimately improving harvests and increasing profitability. The integration of technology, specifically machine learning, represents a significant advancement in farming practices. The survey emphasizes concrete examples and quantifiable outcomes from the literature, contributing novel insights to the existing body of knowledge on crop recommendation systems. By addressing the complexities of crop selection through advanced technologies, this survey aims to bridge the gap between traditional farming practices and modern data-driven approaches.**

***Index Terms*—Agriculture, Crop Recommendation, Machine Learning, Artificial Neural Networks, Sustainability.**

I. Introduction

Agriculture, serving as the backbone of the Indian economy, holds a critical role, utilizing over 60% of the country's land to provide sustenance for a population exceeding 1.3 billion [1]. Soil, the fundamental substrate for crop cultivation, provides essential nutrients, water, oxygen, and structural support. The diverse soil typologies, encompassing alluvial, black, red, and laterite soils, have prompted investigations into the integration of machine learning for crop recommendations based on distinct soil characteristics [2].

In the context of effective agricultural decision-making, particularly in crop selection, this survey addresses the application of machine learning, with a focus on Artificial Neural Networks (ANN), to customize crop recommendation systems according to soil features. The primary objective is to furnish a comprehensive overview of existing methodologies, algorithms, and challenges within this domain, underscoring a commitment to sustainability.

This study endeavors to elucidate the intricacies of agricultural decision-making through a systematic methodology. Utilizing ANN and diverse machine learning algorithms, critical parameters such as N (Nitrogen), P (Phosphorous), K (Potassium) ratios, temperature, humidity, pH, and rainfall are considered. The training of models on a comprehensive dataset, encompassing diverse crops and optimal growing conditions, enables accurate classifications based on specified parameters.

The efficacy of this multi-algorithmic approach, including ANN, in precisely predicting the most suitable crops for cultivation is prominently demonstrated. Through an examination of various studies, the robustness of ANN models alongside other classification algorithms in making accurate predictions is underscored. This synthesis emphasizes the potential of these approaches to assist in optimal crop selection, accounting for soil and environmental conditions. The overarching goal is to enhance crop yields, mitigate risks associated with crop failure, and contribute to sustainable agricultural practices, fostering heightened profitability and efficiency in the agricultural ecosystem [1].

This survey serves as a critical review of existing literature, offering invaluable insights for precision agriculture stakeholders. Meticulously evaluating and comparing the performance of different machine learning algorithms sets the stage for future directions and practical recommendations aimed at enhancing agricultural productivity. Through this comprehensive exploration, the aim is to contribute to the evolution of precision agriculture, bridging the gap between traditional farming practices and modern, data-driven methodologies.

II. Literature Overview

In the dynamic field of agriculture, a transformative shift is underway, characterized by the integration of machine learning algorithms. The exploration commences with pioneering research that highlights decision trees, KNNs, Random Forests, and neural networks as pivotal tools guiding farmers in making informed crop selections [3]. This departure from traditional farming practices signifies a transition into a more data-driven era.

As the survey delves deeper, the focus turns to the crucial role of machine learning algorithms in mitigating crop failure and enhancing productivity, as emphasized by a particular study [1]. Neural networks take center stage, proving instrumental in crop classification. Nevertheless, questions linger regarding the challenges these algorithms face when deployed in the dynamic and unpredictable environment of real-world agriculture.

Another facet of this exploration is presented in a study introducing a production forecasting system intricately linked to farmers through mobile applications [2]. Leveraging GPS, RF algorithms, and boasting a precision rate of 95%, this system paints a promising picture. However, amidst these technological advancements, the discussion falls short of exploring potential hurdles in scaling such systems or adapting them to diverse agricultural contexts.

Further enriching this survey is a contribution that focuses on crop yield prediction through a random forest classifier and a user-friendly website [4]. Serving as a bridge between data analysis and practical decision-making, as the study predicts crop yield based on climate data, broader challenges in agriculture and the potential drawbacks of machine learning models remain unexplored.

A. Technological Evolution

The landscape takes a notable turn as machine learning techniques, particularly Artificial Neural Networks (ANN), assume a central role. The significance of historical data, encompassing vital parameters like temperature, humidity, pH, rainfall, and crop types, is underscored. Despite the enhanced accuracy brought about by random forest algorithms and decision trees, persistent challenges in acquiring and updating historical data, coupled with the system's adaptability to dynamic climate patterns, persist.

Contributing to the discourse, another study focuses on a crop yield recommendation system employing Support Vector Machine (SVM) for precision. The narrative delves into location and crop data, providing insights into nutrient requirements and fertilizer recommendations. However, undisclosed challenges surround the integration of SVM recommendations into existing farming practices.

A comprehensive recommendation system is engineered, considering soil characteristics, types, and crop yield data, with various machine learning algorithms predicting crops based on weather conditions. The exploration aims to guide farmers in selecting suitable crops, yet challenges related to obtaining accurate soil data and the limitations of machine learning algorithms remain elusive.

B. Unexplored Research Challenges

As the exploration deepens, a study implements a crop recommendation system based on soil database information, employing machine learning classifiers. The narrative aims to enhance agricultural productivity, reduce soil degradation, minimize chemical usage, and optimize water resource utilization. However, the challenges of implementing such a system across diverse agricultural landscapes remain unexplored.

Continuing the trajectory, another intelligent system is introduced, incorporating crop suitability and rainfall prediction subsystems. The exploration aims to assist farmers in making informed decisions about crop selection. However, the narrative understates the limitations of relying on predictive models for agricultural decisions and potential challenges in deploying such systems in resource-constrained environments.

As the landscape of crop recommendation systems is surveyed, various machine learning algorithms are reviewed. The exploration emphasizes improved accuracy but leaves the challenges of implementing and maintaining such systems on a large scale shrouded in ambiguity.

The discussion, firmly grounded in research, invites readers to explore uncharted territories where machine learning converges with the soil in a progressive alliance.

III. Methodologies and Approaches

A. Dataset

A commonly used dataset, available freely on Kaggle, was employed in this study [29]. The dataset comprises 2200 rows, each containing information on 8 features: Nitrogen (*N*), Phosphorous (*P*), Potassium (*K*), temperature, humidity, soil pH, rainfall, and the corresponding crop label.

B. Common Approach

Various methodologies were explored by other researchers. When implementing an Artificial Neural Network (ANN) for classification with n inputs and m outputs (depending on the chosen features), the training process utilized the Adam optimizer. This optimizer employs gradient descent to iteratively update the neural network weights, promoting effective learning and adaptation to intricate patterns within the dataset.

Rectified Linear Unit (ReLU) was chosen as the activation function for hidden layers, introducing non-linearity by selecting the maximum value between zero and the output of gradient descent. In the final layer, the softmax function, as described by Equation 1, was applied for classification tasks.

$$\text{Softmax}(x)_i = \frac{e^{x_i}}{\sum_{j=1}^{N} e^{x_j}} \tag{1}$$

Observations during the exploration of Artificial Neural Networks revealed a consistent trend among practitioners: ReLU for hidden layers and softmax for the final layer were consistently applied to introduce non-linearity and facilitate classification. Refer to Figure 1 for a visual representation of the ANN architecture.

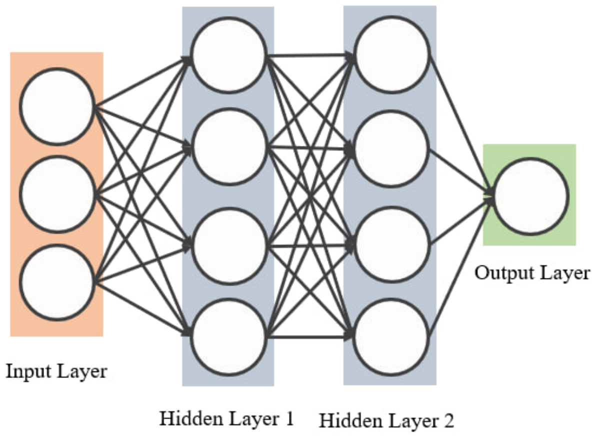

Fig. 1. Neural Network Architecture

The architecture, depicted in Figure 1, represents a standard configuration for neural networks. It includes an input layer, two hidden layers, and an output layer. During the training of our crop recommendation classification model, the input layer processes agricultural features such as soil composition, temperature, humidity, pH values, and rainfall. The hidden layers, strategically positioned between the input and output layers, extract intricate patterns and relationships within the data. The application of activation functions, such as ReLU, in the hidden layers introduces non-linearity, enabling the network to learn complex dependencies. Subsequently, the output layer generates the recommended crop based on the processed information, allowing the neural network to discern intricate relationships in diverse agricultural data.

Researchers also explored various machine learning algorithms, including Decision Trees, Random Forest, and Support Vector Machines, following a shared process outlined in Figure 2. This method involves defining the problem, collecting diverse data, cleaning and transforming it, selecting relevant features, trying different algorithms, and then training, testing, and validating models. The systematic nature of this approach ensures adaptability to complex patterns in datasets and facilitates reliable crop recommendations.

This systematic method extends beyond technicalities to interpreting results and communicating findings effectively. The shared methodology among different algorithms not only aligns diverse approaches but also enhances the credibility and applicability of the results. It provides clear insights into the intricacies of crop recommendation systems, making the study robust and accessible.

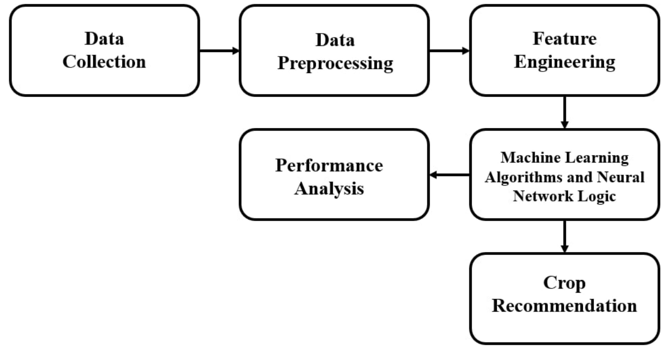

Fig. 2. Methodology Flowchart

IV. Findings and Trends

This study delves into the performance of machine learning algorithms for crop recommendation systems, emphasizing their unique characteristics and considerations. The accuracy levels of different models are summarized in the following table:

A. *Algorithmic Performance*

TABLE I
ALGORITHMIC PERFORMANCE METRICS

Algorithm	Accuracy (%)
Decision Tree (DT) [19]	90
Naïve Bayes (NB) [19]	99
Support Vector Machine (SVM) [19]	10.68
Logistic Regression (LR) [19]	95.22
Random Forest (RF) [26]	97.18
XGBoost [27]	97.4
K-Nearest Neighbors (KNN) [28]	97
Artificial Neural Network (ANN) [25]	98.22

B. *Algorithmic Performance*

Decision Tree (DT): Decision Trees demonstrate a commendable accuracy of 90%, making them suitable for crop recommendation systems. However, their susceptibility to overfitting poses a challenge, particularly in scenarios with intricate relationships between input features and crop recommendations. The rigidity of decision boundaries may hinder their adaptability to diverse agricultural datasets.

Naïve Bayes (NB): Naïve Bayes stands out with an impressive accuracy of 99%. However, its assumption of independence among features may oversimplify complex interactions in agricultural datasets. This oversimplification may lead to suboptimal performance, especially in the presence of intricate, non-linear relationships.

Support Vector Machine (SVM): With an accuracy of 10.68%, Support Vector Machines (SVM) exhibit significant limitations in handling the complexity of agricultural data. SVMs might struggle with non-linear patterns and intricate relationships between environmental factors and crop recommendations, making them less suitable for certain agricultural applications.

Logistic Regression (LR): Logistic Regression achieves a competitive accuracy of 95.22%. However, its limitation lies in capturing non-linear relationships, a crucial factor in agricultural scenarios where crop recommendations often involve intricate and non-linear dependencies on various environmental factors.

Random Forest (RF) and XGBoost: Random Forest and XGBoost showcase impressive accuracy levels of 99% and 99.31%, respectively. While these ensemble methods excel in predictive accuracy, their ensemble nature may pose challenges in interpretability. This lack of transparency could be a drawback, particularly in agricultural applications where interpretability is essential for user trust and acceptance.

Artificial Neural Network (ANN): In agricultural applications, Artificial Neural Networks (ANNs) have shown promising results, achieving an impressive accuracy of 98.22% in crop recommendations [25]. ANNs, known for their adaptability to diverse datasets and effective handling of varying complexities and non-linear relationships, offer valuable insights into intricate and non-uniform data patterns. However, it is essential to acknowledge a trade-off, as ANNs often exhibit reduced interpretability compared to simpler models. This aspect poses a challenge in gaining user trust and acceptance, critical considerations in agricultural decision-making [25].

V. CHALLENGES AND GAPS

A. Data Limitations

One of the primary challenges faced in implementing crop recommendation systems is the availability and quality of data. Insufficient or inaccurate data regarding soil characteristics, weather conditions, and historical crop performance can significantly impact the performance of machine learning algorithms.

B. Interoperability

Integration with existing farming practices and technologies poses a significant challenge. Crop recommendation systems need to seamlessly integrate with diverse agricultural setups, machinery, and data sources. Achieving interoperability is crucial for widespread adoption and practical utility on farms with varying infrastructures.

C. Model Interpretability

While complex machine learning models, such as Artificial Neural Networks, demonstrate high accuracy, their lack of interpretability can be a drawback. Farmers may hesitate to trust and adopt recommendations from models they cannot understand. Future research should focus on developing models that provide clear explanations for their predictions.

D. Environmental Variability

Agricultural conditions are influenced by various environmental factors, including climate change, pest infestations, and soil degradation. The dynamic and unpredictable nature of these factors poses challenges for crop recommendation systems, which need to adapt to changing conditions in real-time.

E. User Acceptance and Education

Farmers' acceptance and understanding of technology play a vital role in the success of crop recommendation systems. Lack of awareness, training, or skepticism among farmers can hinder the adoption of advanced technologies. Bridging the gap between technology developers and end-users is essential for the successful implementation of these systems.

VI. FUTURE RESEARCH DIRECTION

A. Integration of IoT and Sensor Technologies

Future research should explore the integration of Internet of Things (IoT) and sensor technologies to enhance data collection. Deploying sensors in the field to monitor soil moisture, nutrient levels, and crop health in real-time can provide more accurate and timely data for crop recommendation models.

B. Explainable AI for Agriculture

Developing and incorporating explainable AI models tailored for agriculture can address the interpretability challenge. Providing farmers with clear explanations of how the models make recommendations can build trust and facilitate informed decision-making.

C. Personalized and Context-Aware Recommendations

Future research should focus on developing personalized recommendation systems that consider the specific context of each farm. Tailoring recommendations based on the unique characteristics of soil, climate, and historical practices can enhance the relevance and effectiveness of the system.

D. Community-Based Approaches

Exploring community-based approaches to crop recommendation can leverage collective knowledge and experiences. Collaborative platforms that enable farmers to share insights and feedback on crop performance in specific regions can contribute to more robust and context-aware recommendation systems.

VII. CONCLUSION

In summary, this survey explores the landscape of crop recommendation systems, emphasizing the role of machine learning, particularly Artificial Neural Networks (ANN), in revolutionizing traditional agricultural practices. Findings showcase varying accuracy levels among different algorithms, with each demonstrating unique strengths and challenges. Data limitations and interoperability issues emerge as primary hurdles, while the importance of model interpretability is underscored for widespread farmer acceptance.

Future research directions suggest the integration of Internet of Things (IoT) and sensor technologies for improved data collection, the development of explainable AI models, and the exploration of personalized, context-aware recommendations. Community-based approaches also hold potential for leveraging collective knowledge in specific agricultural regions.

In essence, this survey lays a foundation for ongoing research, fostering collaboration between technology developers, researchers, and farmers. As agriculture embraces the digital era, the journey towards optimized crop selection and increased productivity is a collective endeavor promising a more sustainable and prosperous future.

REFERENCES

[1] Priyadharshini A, Swapneel Chakraborty, Aayush Kumar, Omen Rajendra Pooniwala, "Intelligent Crop Recommendation System using Machine Learning", *Proceedings of the Fifth International Conference on Computing Methodologies and Communication (ICCMC 2021).*

[2] Shilpa Mangesh Pande, Prem Kumar Ramesh, Anmol, B R Aishwarya, Karuna Rohilla and Kumar Shaurya, "Crop Recommendation using Machine Learning Approach", *Proceedings of the Fifth International Conference on Computing Methodologies and Communication (ICCMC 2021).*

[3] Doshi, Zeel, Subhash Nadkarni, Rashi Agrawal, and Neepa Shah. "Agro-Consultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms." In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, pp. 1-6. IEEE, 2018.

[4] Mayank Champaneri, Chaitanya Chandvidkar, Darpan Chachpara and Mansing Rathod, "Crop Yield Prediction using Machine Learning".

[5] Kumar, Y. Jeevan Nagendra, V. Spandana, V. S. Vaishnavi, K. Neha, and V. G. R. R. Devi. "Supervised Machine Learning Approach for Crop Yield Prediction in Agriculture Sector." In *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, pp. 736-741. IEEE, 2020.

[6] Suresh, A., P. Ganesh Kumar, and M. Ramalatha. "Prediction of major crop yields of Tamilnadu using K-means and Modified KNN." In *2018 3rd International Conference on Communication and Electronics Systems (ICCES)*, pp. 88-93. IEEE, 2018.

[7] Jejurkar Siddhi, S., S. Bhosale Meghna, and D. N. Wavhal. "Crop Prediction and Diseases Detection Using Machine Learning." (2021).

[8] Rajak, RohitKumar, Ankit Pawar, Mitalee Pendke, Pooja Shinde, Suresh Rathod, and Avinash Devare. "Crop recommendation system to maximize crop yield using machine learning technique." *International Research Journal of Engineering and Technology*, 4, no. 12 (2017): 950-953.

[9] Mahule, Ankit Arun, and A. J. Agrawal. "Hybrid Method for Improving Accuracy of Crop-Type Detection using Machine Learning." *International Journal*, 9, no. 2 (2020).

[10] Nigam, Aruvansh, Saksham Garg, Archit Agrawal, and Parul Agrawal. "Crop yield prediction using machine learning algorithms." In *2019 Fifth International Conference on Image Information Processing (ICIIP)*, pp. 125-130. IEEE, 2019.

[11] Medar, Ramesh, Vijay S. Rajpurohit, and Shweta Shweta. "Crop yield prediction using machine learning techniques." In *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, pp. 1-5. IEEE, 2019.

[12] Jain, Sonal, and Dharavath Ramesh. "Machine Learning convergence for weather-based crop selection." In *2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, pp. 1-6. IEEE, 2020.

[13] Gandge, Yogesh. "A study on various data mining techniques for crop yield prediction." In *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, pp. 420-423. IEEE, 2017.

[14] Akshatha, K. R., and K. S. Shreedhara. "Implementation of machine learning algorithms for crop recommendation using precision agriculture." *International Journal of Research in Engineering, Science and Management (IJRESM)*, 1, no. 6 (2018): 58-60.

[15] Dighe, Deepti, Harshada Joshi, Aishwarya Katkar, Sneha Patil, and Shrikant Kokate. "Survey of Crop Recommendation Systems." (2018).

[16] Jaiswal, Sapna, Tejaswi Kharade, Nikita Kotambe, and Shilpa Shinde. "Collaborative Recommendation System For Agriculture Sector." In *ITM Web of Conferences*, vol. 32. EDP Sciences, 2020.

[17] Ghadge, Rushika, Juilee Kulkarni, Pooja More, Sachee Nene, and R. L. Priya. "Prediction of crop yield using machine learning." *Int. Res. J. Eng. Technol. (IRJET)*, 5 (2018).

[18] Kulkarni, Nidhi H., G. N. Srinivasan, B. M. Sagar, and N. K. Cauvery. "Improving Crop Productivity Through A Crop Recommendation System Using Ensembling Technique." In *2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS)*, pp. 114-119. IEEE, 2018.

[19] Gosai, D., Raval, C., Nayak, R., Jayswal, H., & Patel, A. "Crop Recommendation System using Machine Learning." *International Journal of Scientific Research in Computer Science, Engineering, and Information Technology*, Volume Number(Issue Number), Page Numbers. ISSN: 2456-3307. DOI: https://doi.org/10.32628/CSEIT2173129

[20] Nischitha K, Dhanush Vishwakarma, Mahendra N, Ashwini, Manjuraju M R, "Crop Prediction using Machine Learning Approaches", *International Journal of Engineering Research and Technology*, vol.9 Issue 08, August-2020 ISSN: 2278-0181.

[21] Jayaprakash, S., Nagarajan, M.D., Prado, R.P.D., Subramanian, S. and Divakarachari, P.B., 2021. "A systematic review of energy management strategies for resource allocation in the cloud: Clustering, optimization and machine learning." *Energies*, 14(17), p.5322.

[22] Azizkhan F Pathan, Chetana Prakash, "Attention-based position-aware framework for aspect-based opinion mining using bidirectional long short-term memory," *Journal of King Saud University - Computer and Information Sciences*, 2021, ISSN 1319-1578.

[23] Azizkhan F Pathan, Chetana Prakash, "Unsupervised Aspect Extraction Algorithm for opinion mining using topic modeling," *Global Transitions Proceedings*, Volume 2, Issue 2, 2021, pp. 492-499.

[24] Md Reazul Islam, Khondokar Oliullah, Md Mohsin Kabir, Munzirul Alom, M.F. Mridha, "Machine Learning Enabled IoT System for Soil Nutrients Monitoring and Crop Recommendation," *Journal of Agriculture and Food Research*, Volume 14, 2023, Article 100880, ISSN 2666-1543, https://doi.org/10.1016/j.jafr.2023.100880. (https://www.sciencedirect.com/science/article/pii/S2666154323003873)

[25] Kalaiselvi Bakthavatchalam, Balaguru Karthik, Vijayan Thiruvengadam, Sriram Muthal, Deepa Jose, Ketan Kotecha, Vijayakumar Varadarajan, "IoT Framework for Measurement and Precision Agriculture: Predicting the Crop Using Machine Learning Algorithms," *Technologies*, Volume 10, Issue 1, 2022, Article 13, https://doi.org/10.3390/technologies10010013.

[26] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, Andrey Gulin,"CatBoost: Unbiased Boosting with Categorical Features," In *Advances in Neural Information Processing Systems*, Volume 31, 2018,

[27] S. Rani, S. H. Ahmed, S. C. Shah, "Smart Health: A Novel Paradigm to Control the Chickungunya Virus," *IEEE Internet of Things Journal*, Volume 6, Issue 2, April 2019, pp. 1306-1311, https://doi.org/10.1109/JIOT.2018.2802898.

[28] Vu Khanh Quy, Nguyen Van Hau, Dang Van Anh, Nguyen Minh Quy, Nguyen Tien Ban, Stefania Lanza, Giovanni Randazzo, Anselme Muzirafuti, "IoT-Enabled Smart Agriculture: Architecture, Applications, and Challenges," *Applied Sciences*, Volume 12, Issue 7, 2022, Article 3396, https://www.mdpi.com/2076-3417/12/7/3396.

[29] Atharva Ingle, "Crop Recommendation Dataset," Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset