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 paper explores the application of machine learning, with a focus on Artificial Neural Networks (ANN), in crop recommendation systems tailored to soil features. The overarching vision of this survey paper is to identify the optimal approach for selecting a crop recommendation system, shedding light on its strengths and uncovering drawbacks in alternative algorithms and methodologies. Our objective is to discern the most effective path in crop selection through a critical examination of various recommender systems, contributing to improved decision-making in agriculture.**

***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 [19]. Soil, a foundational component, is indispensable for crop cultivation, furnishing vital nutrients, water, oxygen, and structural support for plant growth. With diverse soil types prevalent in India, ranging from alluvial and black to red and laterite soils, numerous studies have delved into the integration of machine learning techniques to recommend crops based on distinct soil characteristics [20].

In the realm of effective decision-making in agriculture, particularly concerning crop selection, this survey paper explores the application of machine learning, with a specific focus on Artificial Neural Networks (ANN), in crafting crop recommendation systems tailored to soil features. The primary objective is to present a comprehensive overview of existing methodologies, algorithms, and challenges within this domain, with a prominent emphasis on sustainability in agriculture.

Within this context, our research paper aims to unravel the intricacies of decision-making in agriculture by introducing an innovative methodology. Capitalizing on Artificial Neural Networks (ANN) and an array of machine learning algorithms, our approach takes into account critical parameters such as N (Nitrogen), P (Phosphorous), K (Potassium) ratios, temperature, humidity, pH, and rainfall. The training of models on a comprehensive dataset, encompassing diverse crops and their optimal growing conditions, enables accurate classification based on specified parameters.

Our study showcases the effectiveness of this multi-algorithmic approach, including ANN, in precisely predicting the most suitable crops for cultivation. Through an examination of various studies, we underscore the robustness of ANN models alongside other classification algorithms in making accurate and informed predictions. This synthesis emphasizes the potential of these approaches to assist farmers in optimal crop selection, considering soil and environmental conditions. The overarching goal is to enhance crop yields, mitigate risks associated with crop failure, and contribute to sustainable agricultural practices, ultimately fostering increased profitability and efficiency in the agricultural ecosystem [19].

This survey seeks to not only review existing literature but also to contribute valuable insights for researchers, policymakers, and practitioners in the field of precision agriculture. By evaluating and comparing the performance of different machine learning algorithms, we aim to guide future research directions and provide practical recommendations for enhancing agricultural productivity.

II. Literature Overview

In the dynamic field of agriculture, a transformation is underway, marked by the integration of machine learning algorithms. The journey begins with Rashi Agarwal's pioneering research, where decision trees, KNNs, Random Forests, and neural networks emerge as key players in guiding farmers towards informed crop selections [9]. This marks a departure from traditional farming practices towards a more data-driven era.

As the exploration deepens, Priyadharshini A steps into the spotlight, emphasizing the role of machine learning algorithms in mitigating crop failure and boosting productivity [19]. Neural networks take center stage as instrumental in crop classification. However, questions arise about the challenges these algorithms face when applied in the dynamic and unpredictable environment of real-world agriculture.

Shilpa Mangesh Pande introduces a captivating chapter, presenting a production forecasting system intricately connected to farmers through mobile applications [20]. Here, GPS, RF algorithms, and a precision rate of 95

Mayank Champaneri's contribution adds another layer to the exploration, focusing on crop yield prediction through a random forest classifier and a user-friendly website [21]. The narrative bridges the gap between data analysis and practical decision-making. Yet, as the protagonist predicts crop yield based on climate data, broader challenges in agriculture and the potential drawbacks of machine learning models remain unexplored.

A. *The Shift in Technologies*

The exploration takes an interesting turn as machine learning techniques, notably Artificial Neural Networks (ANN), assume a central role. Kumar et al. present a scenario where historical data becomes crucial, featuring vital parameters such as temperature, humidity, pH, rainfall, and crop name [1]. Random forest algorithms and decision trees elevate accuracy, yet the exploration conceals the challenges of obtaining and updating historical data and the system's adaptability to dynamic climate patterns.

Suresh et al. enrich the discussion, focusing on a crop yield recommendation system leveraging Support Vector Machine (SVM) for precision [6]. The narrative zooms in on location and crop data, offering insights into nutrient requirements and fertilizer recommendations. However, the challenges of integrating SVM recommendations into existing farming practices remain shrouded in mystery.

Reddy et al. design a comprehensive recommendation system, considering soil characteristics, types, and crop yield data, with various machine learning algorithms predicting crops based on weather conditions [7]. The exploration aims to guide farmers in selecting suitable crops, yet the unsung challenges of obtaining accurate soil data and the limitations of machine learning algorithms persist in the shadows.

B. *Existing Gaps in Research*

As the exploration deepens, Rajak et al. implement a crop recommendation system based on soil database information, employing machine learning classifiers [8]. The narrative aims to contribute to increased agricultural productivity, reduced soil degradation, minimized chemical usage, and efficient water resource utilization. Yet, beneath the surface, the challenges of implementing such a system across diverse agricultural landscapes linger unexplored.

Continuing the trajectory, Doshi et al. introduce AgroConsultant, an intelligent system incorporating crop suitability and rainfall prediction subsystems [9]. The exploration aims to assist farmers in making informed decisions about crop selection. However, the narrative leaves the limitations of relying on predictive models for agricultural decisions and the potential challenges in deploying such systems in resource-constrained environments understated.

In the denouement of this chapter, as Dighe et al. survey the landscape of crop recommendation systems, they review various machine learning algorithms [10]. The exploration emphasizes improved accuracy but leaves the challenges of implementing and maintaining such systems on a large scale shrouded in shadows.

The discussion, grounded in research, invites readers to explore uncharted territories where machine learning meets the soil in a harmonious dance of progress.

III. Methodologies and Approaches

A. *Modelling of Data*

Modeling data in a survey paper focused on crop recommendation involves a meticulous approach, taking into account various influencing factors. In our research, we utilized Python to develop the model, leveraging essential libraries such as TensorFlow, Keras, NumPy, and Scikit-learn.

For the training phase, we constructed an artificial neural network with n inputs and m outputs, catering to the classification requirements. The training process employed the Adam optimizer, utilizing gradient descent to iteratively update the weights of the neural network. This methodological choice ensures the effective learning and adaptation of the model to intricate patterns within the dataset, contributing to the accuracy and reliability of the crop recommendation system.

For our activation function, we used ReLU for hidden layers, which selects the maximum value between zero and the output of gradient descent, serving as the output of the current neuron. For the final layer, we chose the softmax function which works upon the formula

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

B. *Methodology*

The methodology for this survey paper encompasses the collection of diverse datasets on soil features from official

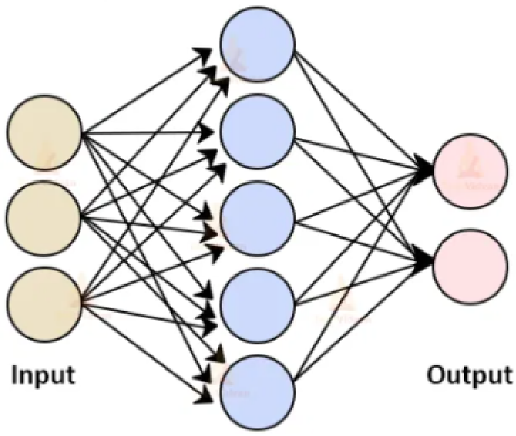

Fig. 1. Neural Network

governmental sources. Key steps in building the data model include:

1. Define the Problem Statement: Clearly articulate the problem statement for the survey paper, focusing on recommending the most suitable crop based on the N-P-K ratio, temperature, humidity, pH value, and rainfall.

2. Gather Data: Collect data from various reliable sources such as research papers, agricultural repositories, and on-farm datasets, ensuring relevance to the defined problem statement.

3. Preprocess the Data: Clean, transform, and prepare the data for analysis by addressing missing values, outliers, normalization, and segmentation into training and testing datasets.

4. Select Appropriate Features: Identify crucial features contributing to crop type prediction, including the N-P-K ratio, temperature, humidity, pH value, and rainfall.

5. Select Suitable Algorithms: Explore various algorithms beyond ANN, considering their applicability. Options may include Decision Trees, Random Forest, Support Vector Machines, and traditional Machine Learning algorithms.

6. Train the Models: Train selected algorithms with the training dataset, adjusting parameters for optimal performance.

7. Test the Models: Evaluate the performance of each algorithm using the testing dataset, assessing accuracy, precision, recall, and F1-score.

8. Validate the Models: Validate selected algorithms on an independent dataset to ensure robust generalization to new data.

9. Interpret the Results: Analyze and interpret the outcomes of different algorithms, identifying the most influential features in predicting crop types.

10. Communicate the Findings: Clearly communicate the survey paper's findings, detailing the methodology, results, and implications for farmers and policymakers.

Acknowledge the versatility of various algorithms beyond ANN in providing valuable insights for crop recommendations.

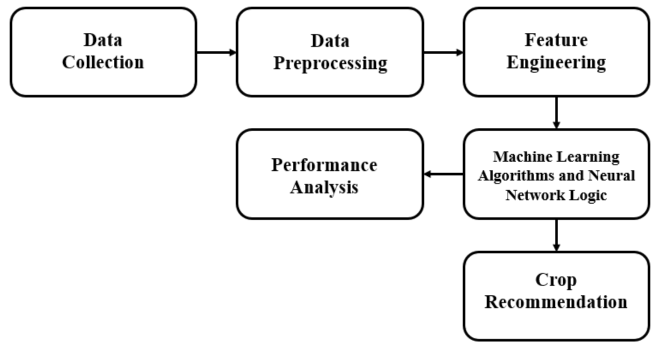

Fig. 2. Flowchart

IV. Findings and Trends

In our survey paper findings, the implemented Artificial Neural Network (ANN) showcased an impressive classification accuracy of 97.67%, underscoring its efficacy in the context of crop recommendation systems. Noteworthy performances were also observed from other compared algorithms, including Naïve Bayes, Random Forest, and XGBoost, all of which exhibited high accuracy levels. While these alternatives displayed competitive results, the ANN model emerged as the preferred choice, primarily due to its exceptional adaptability across diverse datasets.

A. Machine Learning Algorithmic Performance

TABLE I
Machine Learning Algorithmic Performance

Algorithm	Accuracy (%)
Decision Tree (DT)	90
Naïve Bayes (NB)	99
Support Vector Machine	10.68
Logistic Regression (LR)	95.22
Random Forest (RF)	99
XGBoost	99.31

B. Rationale for Choosing ANN:

While the reported accuracy scores indicate strong performance across various algorithms, the following reasons justify the preference for using the Artificial Neural Network (ANN) model in your crop recommendation project:

1. Adaptability to Diverse Datasets: The ANN model excels in adaptability to diverse datasets. In real-world agricultural scenarios, where data may vary in complexity, distribution, and

patterns, ANNs can automatically learn and adjust to different data structures, making them suitable for a wide range of applications.

2. Non-linear Relationships: When dealing with complex and non-linear relationships between input features and crop recommendations, ANNs excel. The ability of ANNs to model intricate patterns can be advantageous in agricultural datasets where interactions between various environmental and crop factors may not follow linear trends.

3. Feature Learning Capability: ANNs possess the capability to automatically learn relevant features from raw data. This is particularly useful in agricultural datasets where the importance of certain features may not be apparent or may change over time. ANNs can adaptively extract meaningful features without manual intervention, potentially improving the model's performance.

4. Robustness to Noise and Variability: The robustness of ANNs to noise and their ability to generalize well can be advantageous in agricultural applications. Agricultural data often contains variability due to changing environmental conditions, and ANNs may handle this variability better than some other algorithms.

5. Parallel Processing and Scalability: For large datasets or future scalability considerations, ANNs, especially deep learning architectures, benefit from parallel processing. This enables efficient training on powerful hardware, facilitating the handling of increased data volumes.

6. End-to-End Learning: ANNs support end-to-end learning, eliminating the need for extensive preprocessing or feature engineering. This can simplify the development pipeline and potentially reduce the manual effort required in preparing the data for modeling.

7. Potential for Improvement with More Data: ANNs, particularly deep learning models, often demonstrate improved performance with larger datasets. If your dataset is expected to grow over time, the ANN model might continue to enhance its performance with more labeled examples.

V. Challenges and Gaps

Despite progress, challenges persist in adopting crop recommendation systems. Insufficient and unreliable data pose hurdles, demanding improved data collection infrastructure. The interpretability of complex models, like ANN, remains a concern for user acceptance. Bridging the gap between sophisticated algorithms and user-friendly interfaces is crucial. Scalability to adapt to dynamic agricultural landscapes and emerging crop varieties requires ongoing efforts from researchers, policymakers, and technologists for robust and transparent solutions.

VI. Future Research Direction

The system has the potential for further enhancement by incorporating the following functionalities:

1) Enhancing the dataset by including a greater number of attributes.
2) Developing a model to distinguish between healthy and diseased crop leaves.
3) Creating a user-friendly website and mobile app for seamless utilization.

VII. Conclusion

This survey explored crop recommendation systems, assessing the performance of machine learning algorithms for optimal crop prediction. The ANN model exhibited a robust classification accuracy of 97.67%, positioning it as the preferred choice for our crop recommendation system. While other algorithms demonstrated high accuracy, the ANN's adaptability and capacity to capture intricate patterns make it well-suited for the complexities of crop prediction.

VIII. References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[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] Doshi, Zeel, Subhash Nadkarni, Rashi Agrawal, and Neepa Shah. "AgroConsultant: 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.

[10] 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).

[11] 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.

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

[13] 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.

[14] 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).

[15] 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.

[16] 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

[17] 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.

[18] 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.

[19] 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).

[20] 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).

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

[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.