2024 International Conference on Computational Intelligence and Computing Applications (ICCICA) | 979-8-3503-0644-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICCICA60014.2024.10584911

Optimizing Crop Predictions: A Machine Learning Approach Incorporating Environmental Parameters and Nutrient Management

Aditi Sabharwal
Computer Science and Engineering
Indira Gandhi Delhi Technical University for Women
New Delhi, India
aditi001phd21@igdtuw.ac.in

SRN Reddy
Computer Science and Engineering
Indira Gandhi Delhi Technical University for Women
New Delhi, India
srnreddy@igdtuw.ac.in

***Abstract*-- In the ever-evolving landscape of agriculture, our research pioneers a revolutionary machine-learning approach to redefining crop recommendations, aligning with the tenets of precision agriculture. We delve into a Kaggle dataset, extracting insights from pivotal parameters like Nitrogen (N), Phosphorus (P), Potassium (K), pH, temperature, humidity, and rainfall. Navigating two scenarios—Scenario 1, comprehensively considering all parameters, and Scenario 2, zeroing in on indispensable NPK nutrients—we deploy Random Forest, Decision Tree, and Support Vector Machine (SVM) models. Meticulously trained and evaluated on an 80:20 dataset split, our research conducts a nuanced comparative analysis, unveiling the intricate interplay between environmental variables and nutrient concentrations. With accuracy as the paramount metric, our findings underscore machine learning's untapped potential in simplifying and elevating crop recommendations, marking a transformative leap forward in precision agriculture. Beyond algorithm intricacies, our study paves the way for sustainable farming, offering indispensable decision support amidst the complexities of modern agriculture. This work encapsulates the transformative role of machine learning in reshaping crop management, steering towards a future characterized by sustainability and efficiency in agriculture.**

Keywords--Precision Agriculture, Machine Learning, Crop Recommendation, Nitrogen, Phosphorus, Potassium (NPK), Environmental Factors, Sustainable Farming

I. INTRODUCTION

Agriculture, the timeless practice of cultivating the land, has significantly transformed through human collaboration and technological advancements [1]. Beyond being a mere food source, it stands as a cornerstone in feeding the global population of billions and is a linchpin in nations' economies. Integrating machinery and technology has revolutionized traditional farming, reducing manual labor and enhancing overall efficiency. However, a pertinent challenge arises with the surge in urbanization: a consequential reduction in available arable land. This doubts our capacity to fulfill agriculture's growing food production needs. Recent studies emphasize the need to increase global food production by over 70% by 2050 to sustain the growing population [1]. Challenges such as diminishing arable land, the necessity for manual labor, and rising capital costs compound the urgency of meeting this demand [2]. This study approaches the analysis from two distinct perspectives. The first encompasses all available parameters, comprehensively understanding the complex interactions shaping nutrient levels. The second strategically exclude specific environmental factors—temperature, humidity, pH, and rainfall—prompting an exploration into the necessity of these variables for accurate nutrient predictions. In recent trends, machine learning models for crop prediction, particularly emphasizing precision farming insights, have been strategically evaluated. Noteworthy models like XGBoost and CBOOST [3-4] are highlighted for their proficiency in achieving accuracy and providing user-friendly recommendations. A paradigm shift is observed in [5-6], introducing IoT integration for real-time soil monitoring. However, inherent gaps persist, necessitating exploration into novel model combinations and an intensified focus on real-time data integration. This study delves into the efficiency of machine learning techniques in predicting crops for precision farming, highlighting the potential for innovative solutions in academic and research realms. Our study strategically evaluates machine learning models for crop prediction, with a specialized focus on precision farming insights. A novel approach is presented, harmoniously amalgamating Random Forest, Decision Tree, and Support Vector Machine models, addressing these voids and aiming for precision and actionable crop predictions in precision agriculture.

This paper delves into the intricacies of modern agriculture, exploring technological advancements showcased in the referenced studies. By leveraging these insights and contributing novel perspectives, our research aims to navigate the complexities of contemporary agriculture, steering toward a more resilient and efficient future. This introductory section lays the foundation for a detailed investigation of the methodology, experimental setup, results, and discussions, presenting nuanced findings with implications for the future of precision agriculture. As we contribute to the discourse on leveraging machine learning in crop recommendations, our insights have broader implications for sustainable and resilient agricultural practices.

II. LITERATURE SURVEY

Thilakarathne et al. [1] introduce a cloud-based platform employing machine learning for precise crop recommendations, advancing precision farming. Ahmed et al. [2] present a soil fertilization nutrient recommendation system employing evolutionary computation techniques that underlines the pivotal role of intelligent technologies in enhancing soil quality and plant nutrition. Shifting focus to India, a formidable agricultural giant, challenges persist despite technological strides. [3].

TABLE I. COMPARATIVE ANALYSIS OF CROP RECOMMENDATION STUDIES

Ref	Crop	Dataset	Parameters	Models	Accuracy (%) (Best)	Outcome
[1]	22 Crops	Crop Recommendation Dataset from Kaggle	NPK, air temperature and humidity, soil pH, and rainfall.	KNN, DT, RF, XGBoost and SVM	RF- 97.18	Precision farming's ML-based crop recommendation platform.
[2]	Cotton, Groundnut, Maize, Rice	Manually Collected	NPK	Improved Genetic Algorithm (IGA)		IGA for crop nutrient optimization
[3]	22 Crops	Manually Collected	NPK, pH, level of the soil, and humidity and rainfall levels	DT, NB, LR, RF, XGBoost	XGBoost- 99	XGBoost improves accuracy, minimizes errors, and optimizes planting efficiency.
[4]	Not Mentioned	Kaggle	NPK, Temperature, Humidity, pH, and rainfall	LR, DT, GNB, MNB, GB, BNB, CNB, SVM, RF, XGBoost, RR, Bagging, SGD, CBOOST	CBOOST- 99.15	Cboost enhances accurate crop recommendations in a user-friendly ML model with chatbot extensions.
[5]	Millet, Groundnut, Pulses, Cotton, Vegetables, Banana, Paddy, Sorghum, Sugarcane, Coriander	Manually Collected and online available dataset is used	Erosion, Water holding, Drainage, Permeability, Soil Color, pH, Texture, Depth.	RF, CHAID, K-NN, NB	The model achieves an 88% prediction accuracy.	Recommends seeds based on soil, boosting productivity and profits for Indian farmers.
[6]	22 Crops	Kaggle and manually collected	NPK, Moisture, Temperature and Humidity, Soil type, Crop type	CatBoost and RF	97.5	Integrates IoT, sensors, and ML for real-time soil monitoring, crop recommendations
[7]	The crops consist of Rice, Wheat, Maize, Bajra, Jowar, and 15 minor crops	Current socio-economic stats & facts about India.	Soil Type, Temperature, Rainfall, Location and Soil Condition	DTK-NN, RF, NN	NN- 91.00	Efficient crop recommendation system, Rainfall Predictor
[8]	Cotton, Groundnut, Maize, Rice	Manually Collected	Temperature, Soil Moisture, pH, NPK	SVM, SVM kernel, DT, and MSVM-DAG-FFO	MSVM-DAG-FFO- 97.3	IoTSNA-CR uses real-time soil data to make precise crop predictions. MSVM-DAG-FFO algorithm achieves high accuracy.
[9]	Bean, Lettuce, Carrot, Cabbage, Beet	Sourced from the Agriculture Department Sri Lanka and other channels.	Temperature, Humidity, Soil pH, Sunlight, and Soil Moisture	NB, SVM	95	Automated crop selection in Sri Lanka with 95%+ accuracy, minimal maintenance
[10]	22 crops	Manually Collected	NPK and pH	K-NN	99.67	NPK sensor-driven system
[11]	Pulse, Rice, Wheat, Potato, Mustard	Manually Collected	NPK	SVM, Adaboost, RF, LR, and SVM.	98	N.P.K-based crop recommendation aids farmers
[12]	Cotton, Sugarcane, Rice, Wheat.	Manually Collected	Soil-specific physical and chemical, Rainfall, and Surface Temperature	RF, NB, and Linear SVM, ET	ET- 99.91	The cropping system enhances productivity for economic gains.

*** Note: Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), K-Nearest Neighbors (KNN), and Extreme Gradient Boosting (XGBoost), Gaussian Naïve Bayes (NB), Logistic Regression (LR), Multinomial naïve Bayes (MNB), Gradient Boosting (GB), Bernoulli naïve Bayes (BNB), Complement Naive Bayes (CNB), SVM, RF, XGBoost, Ridge Regression (RR), Bagging, Stochastic Gradient Descent (SGD), CBOOST, Chi-squared Automatic Interaction Detection (CHAID), K-Nearest Neighbor (K-NN), Naïve Bayes (NB), Neural Network (NN), Multi-Class Support-Vector-Machine-directed acyclic graph-fruit fly optimisation (MSVM-DAG-FFO), Ensembling Technique (ET).*

Factors such as dependence on monsoons, soil quality variations, and unpredictable weather conditions endure. Sharma et al. [3] propose a ground-breaking AI-Farm, a system for recommending crops that employ machine learning to recommend crops appropriate for the prevailing weather conditions and soil composition.

The study introduces an AI-based decision support system for intelligent crop recommendation, addressing challenges Indian farmers face. Leveraging machine learning algorithms on environmental data, the research highlights enhanced precision agriculture. Cat Boosting shows promising accuracy at 99.51%, emphasizing the potential for improved farming practices. The study suggests future enhancements for user-friendliness with chatbot integration [4].

The paper addresses the challenge of suboptimal crop selection by Indian farmers through precision agriculture and an ensemble model-based recommendation system. By employing data mining and machine learning methodologies, the study aims to enhance productivity and contribute to the prosperity of Indian agriculture [5].

This study introduces an Internet of Things (IoT) device empowered by machine learning for accurate soil sensing and crop recommendations, enhancing crop productivity. It uses sensors and modeling techniques to analyze instantaneous soil composition, moisture, and nutrient data. Field experiments validate improved crop management compared to traditional methods, contributing to sustainable agriculture [6].

The study [7] presents AgroConsultant, an intelligent crop recommendation system for Indian farmers, integrating Big Data Analytics and Machine Learning. It aligns with our literature survey on advanced precision agriculture, focusing on sowing season, geographical location, and environmental conditions. Including a Rainfall Predictor and future considerations for predicting crop rotations and economic indicators adds value to the research [7].

The paper introduces IoTSNA-CR, an IoT-enabled crop recommendation system, employing the MSVM-DAG-FFO algorithm for accurate soil nutrient predictions (average 0.969). Accessible through an Android app, the system promotes cost-effective soil management, showcasing the combination of IoT, cloud computing, and machine learning in advancing precision agriculture for enhanced productivity [8].

The paper [9] introduces an IoT-based crop recommendation system for Sri Lankan agriculture, employing Arduino, ML, and AI techniques. It automates crop selection based on environmental factors, ensuring high accuracy (over 95%). The system requires minimal user intervention and self-trains over time, making it suitable for rural and urban areas. The proposed solution stands out as an efficient and cost-effective agriculture consultancy measure.

In this paper [10], the author develops an innovative model for crop prognosis employing NPK sensors, machine learning, and agronomic expertise to offer personalized nutrient recommendations. This data-driven solution promotes precision agriculture, aiming to enhance productivity and sustainability. The system holds the capacity to transform agriculture by addressing the challenges of traditional methods and minimizing environmental impacts.

The paper [11] presents a Crop Suggesting System using machine learning for Bangladesh, achieving 98% accuracy with SVM. It aids farmers in choosing optimal crops based on soil N.P.K. values, promoting efficient land use. The model used is UllahPara, which suggests expansion for broader applicability. Despite limitations, it demonstrates potential for enhancing crop production and supporting agricultural growth.

The paper introduces a highly accurate crop recommendation system utilizing an ensemble of machine learning models like RF, NB, and Linear SVM. With a remarkable 99.91% accuracy, the system assists farmers in making accurate crop choices according to soil characteristics, rainfall, and surface temperature, thereby enhancing overall crop productivity and contributing to economic growth [12].

The study [13] uses machine learning to present a personalized Crop and Fertilizer Recommendation System (CFRS) for Rwanda. The neural network excels in comparative analyses, emphasizing precision and balance. The CFRS aims to optimize agriculture by enhancing yield and reducing environmental impact. Acknowledging limitations, future enhancements include incorporating additional ecological parameters. Precision agriculture, vital for global food security, is emphasized, with the integrated system achieving a commendable 97% accuracy. The study contributes significantly to precision agriculture through decision-support tools combining AI and domain knowledge.

III. Proposed Work

A. *Methodology*

1) *Flowchart*

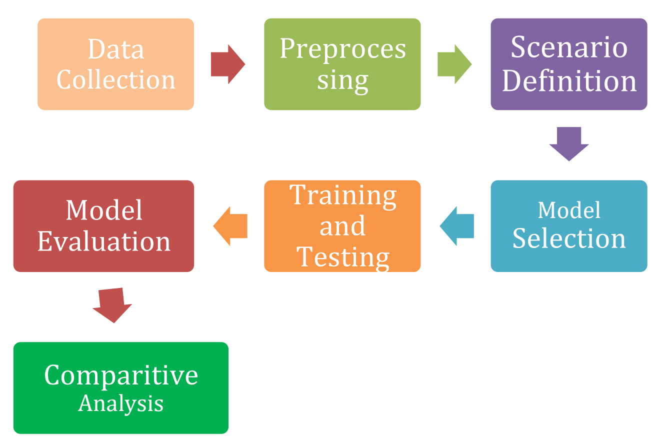

Fig. 1. : ML Model Analysis Workflow

a) Dataset Information:

- *Source: Kaggle crop recommendation dataset [14].*

- *Characteristics:* 2200 instances, 22 different crops.
- *Parameters:* N, P, K, pH, Temperature, Humidity, Rainfall, Crop Name.

b) Scenario Analysis:

- *Scenario 1 (All Parameters): Considered all parameters, i.e., N, P, K, Temperature, Humidity, Moisture, and pH for analysis.*
- *Exploratory Analysis: Understanding dataset size and feature distribution assesses analysis complexity.*
- *Scenario 2 (NPK Parameters): Focused solely on NPK parameters (N, P, K) for crop recommendation.*

c) Models Used:

- *Machine Learning Models:* Employed RF, DT, and SVM.

d) Training-Testing Split Ratio:

- *Data Split:* The dataset underwent an 80:20 split, separating it into training and testing sets.

e) Key Finding:

- *Comparative Analysis:* Investigated the impact of environmental variables and NPK parameters on changes in crop recommendation.
- *Performance Evaluation Measures*: Precision, recall, accuracy, and F1-score metrics provide a holistic understanding of model performance.

f) Study Emphasis:

- *Emphasis:* Stresses the role of multiple parameters in accurate crop recommendations. Highlights environmental factors' significant impact on precision.

2) Models

Our study employed diverse machine-learning algorithms to address the complexities of predicting crop types based on environmental and soil parameters. Strategically selected models include Support Vector Machine (SVM), Random Forest, and Decision Tree, which are known for their robust and accurate predictions.

1) *Random Forest (RF):*

RF was chosen due to its ensemble nature, which combines multiple decision trees. Handles high-dimensional datasets effectively. Minimizes overfitting by combining diverse trees. Improves overall predictive accuracy.

2) *Support Vector Machine (SVM):*

SVM is employed to construct a classification model in multi-dimensional space. Crucial for scenarios with numerous influencing factors (temperature, humidity, pH, nutrient levels). Effectively discerns intricate and non-linear relationships within the dataset. Enhances prediction accuracy in complex agricultural scenarios.

3) *Decision Tree (DT):*

Decision Tree (DT), independently or as part of RF, contributes to decision-making in crop classification. Simplicity and interpretability provide insights into the decision logic. Offers clarity on key parameters influencing crop types. Plays a role in the ensemble approach of RF for improved accuracy.

IV. Results and Implications

Based on the accuracy scores provided in Table II, it is evident that Scenario 1, where all parameters were considered, outperformed Scenario 2, which focused solely on NPK parameters. In Scenario 1, the Random Forest (RF) model exhibited the highest accuracy of 99.31%, followed closely by the Decision Tree (DT) model with an accuracy of 98.63%. The Support Vector Machine (SVM) model achieved an accuracy of 96.13%.

TABLE II. . Accuracy of Models for Different Scenarios

Model	Accuracy of Scenario 1	Accuracy of Scenario 2
RF	99.31	65.90
SVM	96.13	59.77
DT	98.63	62.27

In contrast, Scenario 2 showed lower accuracy scores across all models than Scenario 1. While RF still maintained the highest accuracy among the models, with a score of 65.90%, it was significantly lower than its accuracy in Scenario 1. The DT model performed reasonably well in Scenario 2, achieving an accuracy of 62.27%, while SVM showed the lowest accuracy of 59.77%.

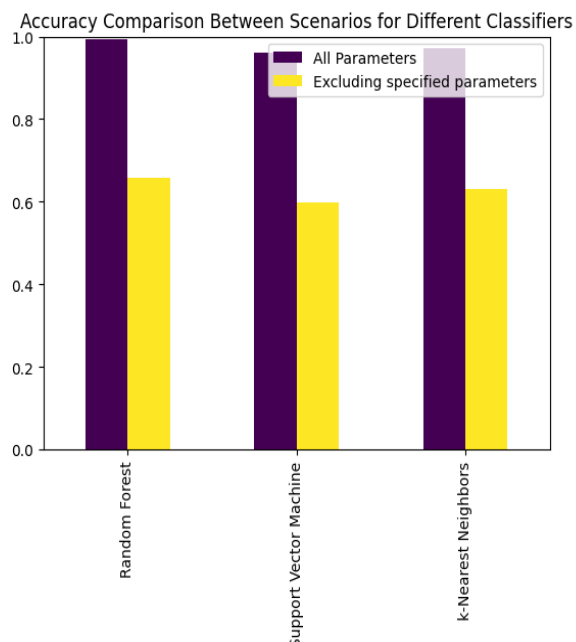

Fig. 2. Showing the Comparative analysis of the accuracy of both Scenarios

This discrepancy in accuracy between the two scenarios underscores the importance of considering all parameters, including environmental factors, in crop prediction models.

Scenario 1, which encompassed a broader set of parameters, demonstrated superior performance, suggesting that incorporating additional environmental factors enhances the precision of crop prediction systems.

Therefore, it is clear that Scenario 1 is the preferred approach for accurate crop predictions, as it provides a more comprehensive understanding of the intricate relationships governing crop classifications. In future implementations, prioritizing the inclusion of all parameters would be beneficial to achieve optimal results in precision agriculture applications.

V. CONCLUSION

In conclusion, our study delved into crop recommendation strategies using machine-learning models, leveraging a dataset sourced from Kaggle containing information on 22 diverse crops and crucial factors such as Nitrogen (N), Phosphorus (P), Potassium (K), pH, temperature, humidity, and rainfall. We meticulously explored two distinctive scenarios to evaluate the efficacy of our models. In Scenario 1, where all parameters i.e., N, P, K, Temperature, Humidity, Moisture, and pH, were considered, a holistic approach was adopted to unravel their collective impact on crop recommendations, excluding the crop name. Conversely, Scenario 2 narrowed the focus to only the vital NPK nutrients. Our analytical framework harnesses sophisticated machine learning models, including Random Forest (RF), Decision Tree (DT), and Support Vector Machine (SVM). Notably, Random Forest emerged as the base model, exhibiting the highest accuracy in Scenario 1, thus underlining its prowess in handling diverse features.

Through an 80:20 split for training and testing, our comparative analysis shed light on the intricate interplay of environmental factors and nutrient concentrations in crop recommendation accuracy. Evaluation metrics quantitatively measured the models' predictive capabilities, particularly accuracy. The outcomes underscored the effectiveness of Random Forest, especially in Scenario 2, suggesting that a concentrated focus on essential nutrient parameters enhances predictive precision.

These findings promise to advance precision agriculture, promote sustainable practices, and empower farmers with intelligent decision support systems. By prioritizing careful parameter consideration and leveraging advanced machine learning techniques, our study contributes to the ongoing efforts to enhance agricultural productivity and sustainability, ultimately benefitting farmers and global food security.

REFERENCE

[1] N. N. Thilakarathne, M. S. A. Bakar, P. E. Abas, and H. Yassin, “A Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming,” Sensors, vol. 22, no. 16, Aug. 2022, doi: 10.3390/s22166299.

[2] U. Ahmed, J. C. W. Lin, G. Srivastava, and Y. Djenouri, “A nutrient recommendation system for soil fertilization based on evolutionary computation,” Comput Electron Agric, vol. 189, Oct. 2021, doi: 10.1016/j.compag.2021.106407.

[3] A. Sharma, M. Bhargava, and A. V. Khanna, “AI-Farm: A crop recommendation system,” in 10th International Conference on Advances in Computing and Communications, ICACC 2021, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/ICACC-202152719.2021.9708104.

[4] S. K. Apat, J. Mishra, K. S. Raju, and N. Padhy, “An Artificial Intelligence-based Crop Recommendation System using Machine Learning,” J Sci Ind Res (India), vol. 82, no. 5, pp. 558–567, May 2023, doi: 10.56042/jsir.v82i05.1092.

[5] S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika, and J. Nisha, “Crop recommendation system for precision agriculture,” 2016 8th International Conference on Advanced Computing, ICoAC 2016, pp. 32–36, Jun. 2017, doi: 10.1109/ICOAC.2017.7951740.

[6] M. R. Islam, K. Oliullah, M. M. Kabir, M. Alom, and M. F. Mridha, “Machine learning enabled IoT system for soil nutrients monitoring and crop recommendation,” J Agric Food Res, vol. 14, Dec. 2023, doi: 10.1016/j.jafr.2023.100880.

[7] Z. Doshi, S. Nadkarni, R. Agrawal, and N. Shah, “AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms,” Proceedings - 2018 4th International Conference on Computing, Communication Control and Automation, ICCUBEA 2018, Jul. 2018, doi: 10.1109/ICCUBEA.2018.8697349.

[8] M. K. Senapaty, A. Ray, and N. Padhy, “IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture,” Computers, vol. 12, no. 3, Mar. 2023, doi: 10.3390/computers12030061.

[9] T. Harinditha Ruchirawya et al., “Crop Recommendation System,” 2020. [Online]. Available: https://www.researchgate.net/publication/346627389

[10] L. Gottemukkala, S. T. R. Jajala, A. Thalari, S. R. Vootkuri, V. Kumar, and G. M. Naidu, “Sustainable Crop Recommendation System Using Soil NPK Sensor,” in E3S Web of Conferences, EDP Sciences, Oct. 2023. doi: 10.1051/e3sconf/202343001100.

[11] S. M. Dipto, A. Iftekher, T. Ghosh, M. T. Reza, and M. A. Alam, “Suitable Crop Suggesting System Based on N.P.K. Values Using Machine Learning Models,” in 2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering, CSDE 2021, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/CSDE53843.2021.9718374.

[12] N. Kulkarni, G. Srinivasan, … B. S.-2018 3rd I., and undefined 2018, “Improving crop productivity through a crop recommendation system using ensembling technique,” ieeexplore.ieee.orgNH Kulkarni, GN Srinivasan, BM Sagar, NK Cauvery2018 3rd International Conference on Computational Systems and, 2018•ieeexplore.ieee.org, Accessed: Jan. 22, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8768790/

[13] M. Zaborowicz, C. Musanase, A. Vodacek, D. Hanyurwimfura, A. Uwitonze, and I. Kabandana, “Data-Driven Analysis and Machine Learning-Based Crop and Fertilizer Recommendation System for Revolutionizing Farming Practices,” mdpi.com, 2023, doi: 10.3390/agriculture13112141.

[14] "Crop Recommendation Dataset." Kaggle. https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset (accessed February 2, 2024).

