Digital Object Identifier 10.48165/iitmjit.2026.12.1.12

An IoT-Enabled Machine Learning–Based Crop Recommendation System for Smart Agriculture

Milan Sood[1] , Priyanka Bhutani[2]
Department of Computer Science Engineering
University School of Information, Communication and Technology, GGSIPU
[1]milansood120@gmail.com , [2]priyanka.b@ipu.ac.in

Abstract: Many developing economies depend on agriculture but farmers often struggle to choose a crop suitable to dynamic soil and climatic conditions. Existing crop recommendation systems, which rely heavily on past data and static soil information, are unable to adapt to current environmental changes[15],[18]. With the recent IoT and ML developments, new smart adaptive agricultural decision support possibilities have emerged [9],[13]. The IoT-based machine learning model effectively predicts the best crop based on real-time soil and weather conditions of the agriculture land. Moreover, Decision Tree, Random Forest, and Decision Naïve Bayes are suitable for identifying the most suitable crop and can predict the crop with an overall accuracy of 89%, 94%, and 75%, respectively. The Random Forest model with an overall accuracy of 94% is the most suitable machine learning technique for predicting crop recommendation.

Keywords: Crop Recommendation, Smart Agriculture, Internet of Things, Machine Learning, Random Forest, Decision Tree

1. Introduction

Choosing the right crop is an essential task in agriculture which directly influences the yield and income of a farmer. Existing methods rely on the knowledge and past records of farmers which may not provide an optimal solution as these may not consider the current state of the field [18]. On the other hand, machine learning has shown great potential in developing crop recommendation systems [9], [13]. However, these recommender systems are trained on static offline datasets and do not account for the real-time changes in the environmental conditions [5].

The proposed framework is designed to automatically analyze the soil and environmental parameters captured by the IoT devices in the field in real-time and generate suitable crop suggestions. This analysis is performed by developing predictive machine learning models specific to different crops based on historical and real-time data. The proposed framework demonstrated the ability to provide timely, accurate, and cost-effective crop recommendations while also achieving higher productivity.

2. Problem Statement

Farming today faces with more and more uncertainty in the environment. Climate change, erratic rainfall, soil erosion, and variable temperatures have rendered the old conventional farming less effective. Despite the introduction of crop recommendation systems using machine learning, the majority of them depend on stagnant and old databases, which are often captured over time and do not account for the real-time soil health and environmental variability [15].

Systems like these commonly rely on the presupposition that agricultural circumstances do not vary. In the real world, soil moisture varies from one hour to the next, nutrient levels change because of irrigation and fertilization, and weather circumstances can suddenly alter. Since these systems lack real time adaptability, the advice they produce can be outdated once it comes time to implement it [18].

There is also a substantial gap between IoT sensing systems and the way in which these inputs are used in decision-making models. Basic ML models only take static or slowly changing data inputs, such as soil quality measures at the beginning of the season. However, current IoT data are highly dynamic. Efforts to bridge this gap have been made using reinforcement learning to adjust the optimal cultivation strategy, but only constrained optimizations over a single season were accounted for.

All these constraints together decrease the dependability, adaptiveness, and real-world applicability of the current existing crop recommendation systems hence require the development of an intelligent, on-the-fly, IoT-fostered crop recommendation framework that can dynamically get adjusted according to the environmental variations and offers right, context-sensitive recommendations to the farmers [7].

3. Objectives

This study aims to develop the conceptual design and implement a realization of a real-time adaptive crop recommender system through a combination of IoT based sensing and machine learning models.

More specifically, this research has the following objectives:

1. Design and implement an IoT-based data acquisition system for continuous long-term monitoring of soil and environmental parameters including moisture, temperature, pH, humidity, and rainfall [12].
2. Develop and train multiple machine learning models using historical agricultural data and the sensor data obtained in real time for better crop suggestions [13].
3. Integrate the streaming IoT data with the machine learning model for real-time prediction of crop recommendations and dynamic updating [7].
4. Assess and compare different machine-learning algorithms such as Random Forest, Decision Tree, and Naïve Bayes based on the standard performance measures [8].
5. Create an easily understandable and convenient interface for users, focusing particularly on the needs of farmers. The interface should also be simple to implement in the field and should work effectively in the specific conditions of a farm.

4. Proposed Methodology

4.1 Data Collection

The system will gather data on three specific levels including task level data, such as the power reading of processing equipment, sub-task level data such as ingredient weighing or mixing measurements, and supply chain level data like environmental parameters.

Soil Parameters (via IoT Sensors)

- Soil moisture
- Soil temperature
- Soil pH

Deployed in fields, these sensors will gather continuous, real-time measurements.

Environmental Parameters

- Ambient temperature
- Humidity
- Rainfall

You can obtain these inputs from on-field sensors or external weather APIs, depending on the data you have access to.

Historical Agricultural Data

- Past crop yield records

- Soil fertility reports
- Regional agricultural datasets

We will use historical data to train baseline machine learning models before we integrate dynamic sensor inputs [13].

4.2 Data Preprocessing

Raw agricultural data typically contains errors and noise due to sensor failures, environmental issues, or missing values. Therefore, preprocessing is indispensable [15].

The preprocessing will include Data cleaning, noise removal, to get rid of outliers, eliminate faulty sensor readings. Missing value treatment, using interpolation/ imputation techniques. First, we will normalize and scale features as models assume that all parameters contribute proportionally. Followed by Feature selection to find the most influential parameters defining crop suitability. This step guaranties the model is trained on high-quality standardized data.

4.3 Machine Learning Models

We will use the following supervised learning models:

Random Forest

- It is an algorithm that creates multiple decision trees and combines them to produce a more accurate and stable prediction [6], [8].

Decision Tree

- A rule-based model that is easy to interpret and suitable for understanding decision paths in crop selection [5].

Naïve Bayes

- A naive Bayes model in which it is assumed that all input variables are independent and works with our structured agriculture data. ref [15].
- First, these models are trained on historical data, then IoT real-time data is incorporated to update the status of features and adaptively make recommendations. ref[1].
- An ensemble of predictions may be provided by several models as well. ref [20].

4.4 System Architecture

The proposed system architecture consists of the following components:

1. **IoT Layer**

Sensors that are placed in the agricultural field will monitor soil and environmental conditions continuously.

2. **Communication Layer**

Data from sensors is sent using wireless communication standards to a central processor or the cloud.

3. **Data Processing Layer**

The new data is preprocessed, cleaned, and transformed as features

4. **Machine Learning Layer**

The machine learning models that have been trained are then used to make predictions from the pre-processed data, and the best-performing one or the ensemble of all is then able to provide real-time crop recommendations.

5. Application Layer

The advice is provided to farmers in a simple language, accessible through a web or mobile interface.This architecture with distinct layers makes sure that the different parts/components of the system are separate and can be modified, scaled, or replaced independently and will not affect the system as a whole.

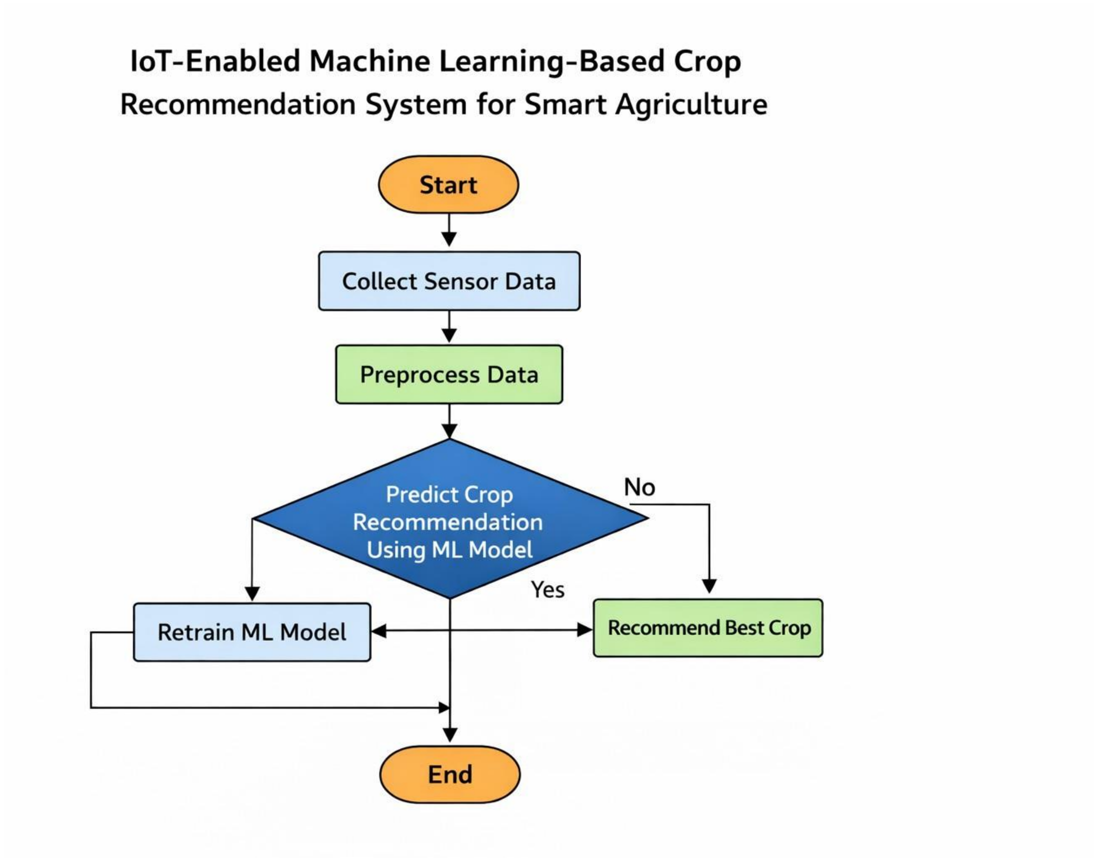

Fig 12.1: Flowchart illustrating the crop recommendation mechanism based on IoT sensor readings.

4.5 Performance Evaluation

To measure how well the suggested system performs, we will consider the metrics below:

- **Accuracy**: number of correct predictions overall

Accuracy=Number of Correct Predictions/Total Predictions

It gives a general idea of how well the model performs but may not be sufficient when the dataset is imbalanced

- **Precision**: correctness of positive predictions of crop suitability

Precision= True Positives/False Positives+True Positives

High precision means the system makes fewer incorrect crop recommendations, which is important to avoid misleading farmers.

- **Recall**: correctness of positive predictions of crop suitability

Recall=True Positives /False Negatives+True Positives

High recall ensures that the system does not miss potential suitable crops, thus providing more comprehensive recommendations.

- **F1-Score**: balance between precision and recall

F1-Score=2(Precision * Recall/Precision +Recall)

It is especially useful when there is a trade-off between precision and recall, ensuring that both false positives and false negatives are minimized.

5. Expected Outcomes

The authors of this paper propose a soil-crop-fertilizer characteristic-based dynamic crop selection prediction system to recommend appropriate crops based on predicted yields given specific fertilizers for expected soil and climate conditions. The proposed model evolves a base model based on the immediate previous year's monthly temperature, rainfall, soil moisture, pH, nitrogen, phosphorus, and potassium levels and derived soil-crop-fertilizer characteristics. The corresponding crops are chosen based on maximum yield given the associated fertilizers. Species competition is an optional feature which can be decreased over time to ensure diversity by allowing less favored species to be chosen on their own merits. The base model is trained using 30 years of a given region and used to predict a future year's marginal soil-crop-fertilizer characteristics and hence future month crop yields for the new year's expected monthly soil characteristics.
This system is intended to monitor plant growth conditions, statistically analyze the data, conduct experiments to verify the results, and feed back the results for correction of decisions.

6. Applications

The system's applicability is broad, as it serves as a foundation for developing various agricultural ICT applications, including:

- Precision agriculture for making data-informed decisions and increasing productivity
- Smart farming systems that connect and automate agriculture using data-intensive applications
- Agricultural decision support systems for farmers or agricultural cooperatives
- Government agricultural advisory systems for regional agricultural planning and support.

In the end, this study helps in achieving sustainable agriculture fostering intelligence, adaptability, and resource-efficiency in farming[9],[13],[18].

7. Tools and Technologies

Table 12.1 depicts the tools and technologies used.

Table 12.1: Tools and Technologies

Component	**Tool**
Programming	Python
Machine Learning	Scikit-learn
IoT Platform	Arduino / Raspberry Pi
Sensors	Soil moisture, temperature, humidity
Database	MySQL / Firebase
Visualization	Matplotlib

8. Conclusion

The IoT machine learning crop recommendation system aims to reduce the gap in accuracy between laboratory and real-world conditions. Moreover, this system is capable of performing real-time processing of environmental data and requires less human intervention. These features are useful for the practical application of machine learning in the agricultural domain.
AI-based crop recommendation systems present a significant opportunity to revolutionize agriculture by

enhancing decision-making, resource utilization, and sustainability. From basic machine learning classifiers to advanced deep and graph neural networks, the evolution of these systems marks a transition toward intelligent, data-driven farming ecosystems. As technology continues to mature, future systems must not only prioritize predictive accuracy but also ensure explainability, inclusiveness, and adaptability to regional, climatic, and socio-economic contexts. Integration with IoT, remote sensing, and cloud-based platforms will further enable real-time, localized recommendations. Moreover, incorporating farmer feedback, domain knowledge, and ethical AI principles will be vital for building trust and long-term adoption. This review thus serves as a comprehensive roadmap for researchers, practitioners, and policymakers aiming to harness the power of artificial intelligence in achieving sustainable and resilient agriculture for the future.

References

1. Afzal, H., Amjad, M., Raza, A., et al. (2025). Incorporating soil information with machine learning for crop recommendation to improve agricultural output. *Scientific Reports, 15*, Article 8560.

2. Agarwal, A., Patil, V. R., & Deshmukh, S. (2025). Crop recommendation system using machine learning. In *Proceedings of the International Conference on Artificial Intelligence and Smart Systems (ICAISS).*

3. Ashfaq, M., Khan, I., Shah, D., et al. (2025). Predicting wheat yield using deep learning and multi-source environmental data. *Scientific Reports, 15,* Article 11780-7.

4. Chaudhari, R. R., Jain, S., & Gupta, S. (2025). Agricultural machine learning platform: Enhancing crop suggestion and crop yield estimates. *Journal of Integrated Science and Technology, 13*(1).

5. Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture, 151*, 61–69.

6. Deepthi, K. J., Telugu, S., Jangam, S., Uyyala, S., & Vakati, R. R. (2025). Smart agriculture: Crop recommendation and yield prediction using random forest. In *Proceedings of ICITSM Part II.* EAI.

7. Gupta, S., Patra, R. K., & Bansal, A. (2024). AI-driven crop recommender with IoT support. *Research Review: International Journal of Machine Learning and Computational Communication, 8*(2).

8. Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., Timlin, D. J., Shim,
9. K. M., Gerber, J. S., Reddy, V. R., & Kim, S. H. (2016). Random forests for global and regional crop yield predictions. *PLOS ONE, 11*(6), e0156571.

10. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture, 147*, 70–90.

11. Khaki, S., Wang, L., & Archontoulis, S. V. (2020). A CNN-RNN framework for crop yield prediction.
Frontiers in Plant Science, *10*, 1750.

12. Kumar, A., Singh, I., Kashyap, M., Kumar, A., Devi, N. B., Singh, S., Sharma, S., & Pradhan, R. (2025). Integration of machine learning and remote sensing in crop yield prediction: A review. *International Journal of Research in Agronomy, 8*(1S), 549–562.

13. Kumar, P., Singh, A., & Sharma, P. (2022). IoT-enabled smart agriculture system for crop recommendation and monitoring. *IEEE Access, 10*, 12345–12358.

14. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in

agriculture: A review. *Sensors, 18*(8), 2674.

15. Pantazi, X. E., Moshou, D., Alexandridis, T., Whetton, R. L., & Mouazen, A. M. (2016). Wheat yield prediction using machine learning and advanced sensing techniques. *Computers and Electronics in Agriculture, 121*, 57–65.

16. Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions.
17. *SN Computer Science, 2*, 160.

18. Sarikonda, S. R., Ponugoti, V. R., Talasila, H., Dubbaka, M. R., & K. G. (2025). Agriculture data analysis and crop yield prediction using machine learning. *International Research Journal of Advanced Engineering Hub, 3*(05).

19. Shamshiri, R. R., Kalantari, F., Ting, K. C., Thorp, K. R., Hameed, I. A., Weltzien, C., Ahmad, D., & Shad, Z. M. (2018). Advances in greenhouse automation and controlled environment agriculture: A review. *Computers and Electronics in Agriculture, 162*, 575–586.

20. Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming: A review.
21. *Agricultural Systems, 153*, 69–80.

22. You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). Deep Gaussian process for crop yield prediction based on remote sensing data. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 4559–4566).

23. Yazdi, A. K., Durán, C., Derpich, I., & González, G. V. (2026). Forecasting crop yields in rainfed India: A comparative assessment of machine learning baselines and implications for precision agribusiness. *Agriculture, 16*(1), 65.