Weather-Based Crop Recommendation Using Machine Learning Survey

Shrushti Waghmare[1] Nisha Taware[2] Rohit More[3] Ganesh Ranjane[4] Parth Kank[5]
[1,2,3,4,5]Department of Artificial Intelligence and Machine Learning
[1,2,3,4,5]Zeal Polytechnic, Pune, Maharashtra, India

Abstract — Crop selection depends on soil nutrients, soil pH, temperature, humidity, rainfall, season, and local growing conditions. Machine-learning methods can combine these variables to recommend crops from historical agricultural data. This survey reviews machine-learning approaches used for crop recommendation and related agricultural prediction, including classifier-based systems, soil and weather data integration, and sensor-supported decision tools. Random Forest and boosting methods are promising candidates, but performance on a static dataset does not guarantee reliable results in a different region or season. The paper identifies practical considerations involving regional validation, rainfall variability, interpretability, and input-data quality. It then describes a proposed Weather-Based Crop Recommendation System with rainfall-scenario comparison, understandable soil guidance, and image-quality checks for uploaded land or soil photographs. Photographs are intended for quality screening only and cannot reliably replace measured values for nitrogen, phosphorus, potassium, or pH. The proposed system remains a design for future implementation and has not yet been experimentally validated.

Keywords: Crop Recommendation, Machine Learning, Random Forest, Soil Nutrients, Weather, Rainfall Risk, Precision Agriculture

I. Introduction

Agriculture is influenced by changing weather and differences in soil properties. Choosing a suitable crop requires considering multiple factors together rather than relying on a single measurement. Common crop-recommendation datasets include nitrogen (N), phosphorus (P), potassium (K), pH, temperature, humidity, rainfall, and a crop label. A classifier can learn patterns between these inputs and known crop labels, then predict a likely crop for new input values.

Machine learning is also used for crop-yield prediction, soil analysis, disease detection, and broader agricultural decision support [1], [5], [6]. These tasks are related but not identical. Crop recommendation suggests a crop class under supplied conditions, whereas yield prediction estimates the amount of production. A model should therefore be evaluated according to its intended task and the data on which it will be used.

For practical use, a recommendation tool should communicate its limitations and avoid treating a model output as a guarantee. Rainfall may differ from historical averages, soil measurements may be incomplete, and a model trained in one region may not transfer well to another. This survey summarizes relevant research directions and uses them to motivate a student project that emphasizes scenario comparison, transparent guidance, and input quality.

II. Objectives and Survey Method

The objectives are to: (1) review machine-learning approaches used for crop recommendation and related agricultural prediction; (2) compare common input variables and modelling approaches; (3) identify limitations related to regional generalization, rainfall variation, and interpretability; and (4) outline a proposed system with crop prediction, soil guidance, rainfall scenarios, and photograph-quality screening.

This paper is a focused narrative survey based on accessible scholarly article records and review publications. The selected sources address crop-yield prediction, soil- and weather-based crop recommendation, explainable classification, and data collection through sensors. This is not an exhaustive systematic review and does not claim formal database screening or PRISMA procedures.

A. Machine Learning for Agricultural Prediction

Van Klompenburg, Kassahun, and Catal reviewed machine-learning methods for crop-yield prediction and found that weather and other environmental variables are commonly used across the literature [1]. Their review provides useful context for selecting input features and evaluating agricultural models. However, yield prediction differs from crop recommendation, so conclusions about one task should not be directly treated as proof of performance for the other.

B. Soil and Weather-Based Crop Recommendation

Recent crop-recommendation studies combine soil nutrient measurements with weather variables and apply classification algorithms such as Random Forest [2]. The common feature set includes N, P, K, pH, temperature, humidity, and rainfall. Such systems can produce a crop label from structured input data, but their reliability depends on dataset quality, label correctness, and whether the training data represent the target location and season.

C. Classifier Comparison and Explainability

Comparative studies examine different classifiers for crop and soil-based recommendation and discuss interpretability and deployment challenges [3]. A model should not be selected only because it has a high accuracy score on one test split. Precision, recall, F1-score, a confusion matrix, and tests on independent or regionally distinct data can provide a fuller view. Explanations such as feature importance can describe which variables influence predictions, but they do not independently establish that a soil nutrient is deficient.

D. Sensor-Driven Agricultural Data

Sensor-supported approaches aim to collect more current soil and weather observations and combine them with machine-learning models [4]. IoT-based collection can improve the timeliness of input data, but deployment depends on sensor calibration, maintenance, connectivity, cost, and coverage. Sensor values also need validation; current readings should

All rights reserved by www.ijsrd.com

not be assumed to be accurate without checking the measurement process.

E. Findings from Review Literature

Review studies describe the use of methods such as Random Forest, XGBoost, Support Vector Machines, Decision Trees, and k-nearest neighbours in agricultural prediction [1], [5], [6]. Across this work, important recurring considerations include data representativeness, model evaluation, integration of environmental variables, and the gap between benchmark results and field use. These findings support comparing a simple baseline with candidate models and testing beyond a single random split.

III. COMPARATIVE SUMMARY

The reviewed approaches differ in task and data. Crop-yield reviews summarize environmental features and prediction methods [1], while crop-recommendation studies use soil and weather variables to predict crop labels [2], [3]. Sensor-based systems seek current measurements but must address calibration and deployment constraints [4]. Recent reviews identify data representativeness, interpretability, and validation as recurring concerns [5], [6].

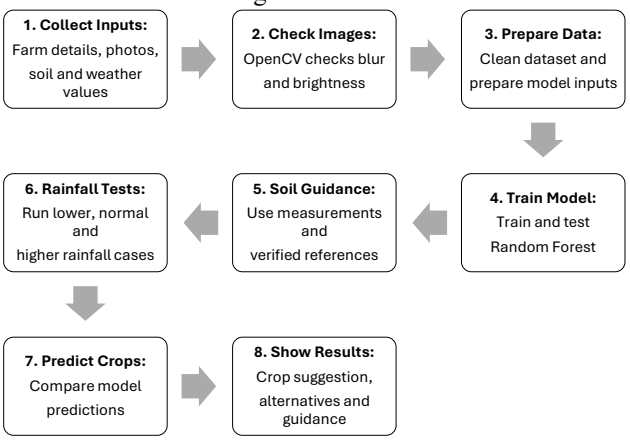
Research direction	Contribution	Consideration
ML yield review [1]	Summarizes agricultural ML methods and features	Yield prediction differs from crop recommendation
Soil/weather crop recommendation [2]	Uses nutrient and climate inputs	Local and seasonal validation needed
Explainable classifiers [3]	Model comparison and interpretation	Importance is not nutrient diagnosis
Sensor-based systems [4]	Integrates measured data	Calibration and access constraints
Review literature [5], [6]	Summarizes trends and challenges	Static data may not represent farms

IV. RESEARCH GAP

The literature establishes crop recommendation from soil and weather inputs as an existing research direction. Practical opportunities for the proposed project include comparing recommendations under lower, normal, and higher rainfall scenarios; presenting understandable soil guidance; screening uploaded photographs for blur and poor brightness; and clearly communicating model limitations. These are proposed design goals rather than a claim that no previous system has used similar features. A broader literature search and direct benchmarking would be needed to establish novelty.

V. PROPOSED SYSTEM

The planned system will accept measured soil and weather parameters and use a trained classifier, initially Random Forest, to recommend a suitable crop. Inputs may include N, P, K, pH when available, temperature, humidity, and rainfall. Data preprocessing, duplicate checks, label checks, and a held-out test set will be used. Decision Tree, k-nearest neighbours, or Gaussian Naive Bayes may be evaluated as baselines before selecting the final model.



Planned tools include Python, Pandas, NumPy, Scikit-learn, OpenCV, and Streamlit. Evaluation will consider accuracy, precision, recall, F1-score, and a confusion matrix, with independent or regionally distinct testing where feasible. Since the system has not yet been implemented, this survey reports no measured accuracy or field results.

VI. CONCLUSION

The surveyed literature shows that machine learning can support agricultural decision-making and that soil nutrients and weather variables are frequently used in crop-related prediction. Random Forest and boosting-based methods are reasonable candidates for comparison, but performance depends on data quality, local relevance, and evaluation design. The proposed student project will build on established crop recommendation by exploring rainfall-scenario comparison, clear soil guidance, and photograph-quality screening. Future work should implement the prototype, validate it with reliable data, compare baseline algorithms, and seek expert feedback before using its recommendations in real farming.

REFERENCES

[1] T. van Klompenburg, A. Kassahun, and C. Catal, "Crop yield predictionusing machine learning: An extensive and systematic literature review," Computers and Electronics in Agriculture, vol. 177, Art. no. 105709, 2020, doi: 10.1016/j.compag.2020.105709.
[2] S. Thevar, S. Shaikh, A. Mujawar, and P. Rastogi, "Crop RecommendationBased on Soil and Weather Data," International Journal of Creative Research Thoughts, vol. 14, no. 4, 2026, doi: 10.56975/ijcrt.v14i4.305278.
[3] H. Trivedi, H. Chauhan, S. Patel, M. N. Patel, P. Patel, M. Parmar, and C.Dabhi, "Explainable Crop and Soil-Based Recommendation System for Next-Generation Smart Agriculture Using Model Classification," Pertanika Journal of Science & Technology, vol. 34, no. 4, 2026, doi: 10.47836/pjst.34.4.08.
[4] V. Balaji K., K. S. Balamurugan, T. Thanarajan, and A. M. S. Sundarapandi, "Machine Learning-Based Crop Recommendation System Integrating Soil Properties and

All rights reserved by *www.ijsrd.com*

Weather Conditions with IoT-Driven Data Collection," Agrociencia, early access, 2026, doi: 10.47163/agrociencia.v60i1.3428.
[5] "Crop recommendation in precision agriculture: a systematic literaturereview of methods, trends, and challenges," African Scientific Reports, 2026.
[6] "Crop yield prediction using machine learning: A systematic literaturereview," Agricultural Technology, 2024, doi: 10.1016/j.atech.2024.100718.

All rights reserved by *www.ijsrd.com*