IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

SMART CROP RECOMMENDATION SYSTEM

Dr. Beenarani Manoj[1], Manish Kumar Mandal[2], Sreejith S[3], Sanjith Harihar T[4]

Faculty Of Engineering and Technology, Department of Computer Science and Engineering,

SRM Institute of Science and Technology, Vadapalani, Chennai, Tamil Nadu, India[1,2,3,4]

Abstract: Choosing the right crop for a field strongly affects yield, resource use and the economic stability of rural households. Conventional expert advice and generalised advisories seldom reach field-level precision, and many digital tools either return a prediction with no explanation or mix model outputs with plain statistical records. This paper presents Smart Crop Recommendation, a web application that keeps the two apart: a machine learning recommendation engine and a regional statistical lookup. The recommendation engine is a Random Forest classifier of 100 decision trees, trained on a 2,200-sample dataset covering 22 crops. It takes seven soil and weather inputs: Nitrogen, Phosphorus, Potassium, pH, temperature, humidity and rainfall. Rather than a single unexplained answer, it ranks every crop and returns the top three with their class probabilities and a comparison of the entered values against recorded values. The lookup component gives a district-level profile for 38 districts of Tamil Nadu, taken from the official statistical handbook (2021–22), and involves no machine learning. On a held-out test set of 440 samples (20% of the data), the model reached an accuracy of 99.32% and a top-3 hit rate of 100.00%. The system also validates user inputs and supports United Nations Sustainable Development Goal 2 (Zero Hunger).

Keywords: Precision Agriculture, Crop Recommendation, Random Forest, Explainable AI, Machine Learning, Agro-Climatic Profile, Sustainable Development Goal 2.

1. INTRODUCTION

Crop selection is among the most important decisions in precision agriculture. A crop that suits the local soil nutrients and climate reduces yield loss, cuts down fertilizer and water waste, and shields farmers from serious financial stress. Soil fertility depends mainly on the macronutrients Nitrogen (N), Phosphorus (P) and Potassium (K), together with soil pH (acidity or alkalinity), while temperature, relative humidity and rainfall shape the plant growth cycle.

Machine learning classifiers handle tabular data well, yet the crop advice systems available today share several shortcomings:
Opaque output: a single predicted crop is returned with no confidence score or supporting comparison.
Mixed data sources: speculative model output is shown alongside fixed regional crop statistics without saying which is which.
No boundary checks: users are not warned when their inputs fall outside the range the model was trained on.

To address these gaps, we developed Smart Crop Recommendation, a dual-pipeline web application that keeps machine learning inference and statistical lookup strictly separate. The first pipeline uses a Random Forest model to predict the three most suitable crops from seven laboratory-tested parameters. The second presents official district-level cultivated-area records for crops in Tamil Nadu. Because the application reports class probabilities, feature comparisons and input range warnings, users can see how each result was produced.

2. LITERATURE REVIEW

Moving from experience-based farming to data-driven Decision Support Systems (DSS) is a central theme in modern agriculture. Many studies have applied machine learning to crop selection, yield prediction, soil management and overall productivity.

Traditional and Rule-Based Methods: Early advisory systems depended on fixed lookup tables and expert-written rules, for example conditional checks on soil pH. They are easy to read and interpret, but they cope poorly with many interacting variables and with the non-linear relationships between soil and environmental parameters.

Machine Learning in Agriculture: Machine learning is now widely used to support agricultural decisions. The reviews by Liakos et al. [4] and van Klompenburg et al. [9] describe how supervised learning has been applied to yield prediction, soil management and crop selection, with Support Vector Machines (SVM), k-Nearest Neighbors (k-NN),

 This work is licensed under a Creative Commons Attribution 4.0 International License

Decision Trees and Naïve Bayes among the most frequently tested algorithms. Chlingaryan et al. [2] reviewed work on yield prediction and nitrogen status estimation in precision agriculture.

Ensemble Models and Random Forest: Random Forest, introduced by Breiman [1], builds many decision trees using bootstrap aggregation (bagging) and random feature selection, and combines their votes. Averaging over trees lowers variance, stabilises predictions and lets the model capture non-linear relationships between inputs. Pudumalar et al. [5] applied an ensemble approach to crop recommendation from soil-related features and showed that it suits precision agriculture.

Explainability and Trust: Farmers are more likely to trust a machine learning tool when they can see why it made a recommendation, which is where Explainable Artificial Intelligence (XAI) helps. Local Interpretable Model-Agnostic Explanations (LIME) by Ribeiro et al. [6] and SHapley Additive exPlanations (SHAP) by Lundberg and Lee [7] explain individual predictions, while feature importance scores show which factors drive a recommendation most. Official statistics such as the Tamil Nadu Statistical Handbook [3] add regional context for planning. Work of this kind also contributes to United Nations Sustainable Development Goal 2 (Zero Hunger) [10] through more sustainable farming and better food security.

3. PROPOSED SYSTEM ARCHITECTURE

The proposed architecture is divided into separate modules for data acquisition, AI prediction, regional lookup and presentation.

3.1 System Features

Pipeline 1 (Soil Analysis AI): Takes 7 soil and weather inputs, applies feature scaling, runs them through the trained Random Forest model and returns the top 3 crops with probability scores.

Pipeline 2 (Tamil Nadu Regional Profile): Reads the published state agricultural statistics (2021–22) for 38 districts and lists crops by actual cultivated area in hectares.

Shared Crop Library: Holds crop cards with images, source tags (AI Model, TN Data, or both) and categories, so that users can place results in context.

4 METHODOLOGY

4.1 Dataset and Features

The model uses a crop recommendation dataset from Kaggle with 2,200 samples spread over 22 crop classes. Each sample has seven numeric attributes:

- Nitrogen: available soil nitrogen, measured in kilograms per hectare.
- Phosphorus: available soil phosphorus, measured in kilograms per hectare.
- Potassium: available soil potassium, measured in kilograms per hectare.
- Temperature: average seasonal temperature, in degrees Celsius.
- Humidity: average seasonal relative humidity, as a percentage.
- pH: soil acidity or alkalinity on a scale of 0 to 14.
- Rainfall: seasonal rainfall, in millimeters.

4.2 Data Preprocessing

Ordering and Validation: Input features are arranged in a fixed order: Nitrogen, Phosphorus, Potassium, temperature, humidity, pH and rainfall. Values outside the training ranges trigger an extrapolation alert; for example, rainfall outside 20–299 mm.

Transformation: Features are scaled first with a min-max scaler and then with a standard scaler.

 This work is licensed under a Creative Commons Attribution 4.0 International License

4.1 Model Configuration and Training

A Random Forest classifier with 100 estimators was trained on a stratified 80:20 split, giving 1,760 training samples and 440 test samples. The model was built with the Scikit-learn library [8].

5 EXPERIMENTAL RESULTS

5.1 Model Evaluation Metrics

The model was evaluated on the 440 held-out samples using classification accuracy, weighted precision, recall and F1-score, and the top-three hit rate.

- Overall accuracy: 99.32%.
- Weighted precision: 99.37%.
- Weighted recall: 99.32%.
- Weighted F1-score: 99.32%.
- Correct classifications: 437 out of 440.
- Top-three hit rate: 100%, covering all 440 test samples.

5.2 Error Analysis and Confusion Matrix

Only 3 of the 440 test samples were misclassified at the top-one level:

- Two rice samples were predicted as jute.
- One moth bean sample was predicted as lentil.

In every misclassified case the true crop still appeared among the top three predictions, which is why the top-three hit rate is 100%.

All three errors occurred between crops with strongly overlapping nutrient profiles and climatic ranges.

6 DISCUSSION

Operational Transparency and Trust:
Every output is labelled either as a machine learning prediction or as an official statistical record, so the two cannot be confused. Rainfall extrapolation warnings and the requirement for laboratory-tested soil data add further safeguards and make the results easier to trust.

Practical Usability:
The interface gives plain-language hints and units for each input field. Showing the three most probable crops instead of one fixed answer gives farmers realistic alternatives.

Observed Limitations:
The model covers only 22 crop classes, so regional staples such as castor and cholam can be explored only through the Tamil Nadu statistical lookup. In addition, the training dataset is not specific to Tamil Nadu, boundary geometry constraints meant that searchable dropdown menus were used instead of maps, and some crop entries still use placeholder illustrations.

7 CONCLUSION AND FUTURE WORK

The Smart Crop Recommendation System takes a dual-pipeline approach to precision farming. A Random Forest model with 99.32% accuracy, combined with district-level statistical lookup, gives farmers and researchers guidance that is both usable and transparent.

Future Scope:

Regional Soil Training:
Collect location-specific soil and climate data for Tamil Nadu and retrain the model, possibly with gradient boosting frameworks such as XGBoost or LightGBM.

Automated Weather Application Programming Interface:
Fetch real-time climate parameters automatically using the user's GPS location.

Advanced Explainability:
Add SHAP and LIME to give detailed local feature attribution for each prediction.

Multilingual Support:
Add local language interfaces such as Tamil, Telugu and Hindi to improve accessibility.

REFERENCES

[1] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001
[2] A. Chlingaryan, S. Sukkarieh, and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review," Computers and Electronics in Agriculture, vol. 151, pp. 61–69, 2018
[3] Government of Tamil Nadu, Statistical Handbook 2021–22, Agriculture Department, Tables 4.1 and 4.4
[4] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis,"Machine Learning in Agriculture: A Review," Sensors, vol. 18, no. 8, p. 2674, 2018
[5] S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika, and J. Nisha, "Crop recommendation system for precision agriculture," in Proceedings of the 8th International Conference on Advanced Computing (ICOAC), IEEE, pp. 32–36, 2016
[6] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You?: Explaining the Predictions of Any Classifier," in Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135–1144, 2016
[7] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, pp. 4765–4774, 2017
[8] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011
[9] T. van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: A systematic literature review," Computers and Electronics in Agriculture, vol. 177, p. 105709, 2020
[10] United Nations, "Sustainable Development Goal 2: Zero Hunger, "UN Department of Economic and Social Affairs