International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 3, pp.194-201, March 2026.
www.ijersem.com eISSN- 3107-9075

Crop Prediction Based on Characteristics of Agricultural Environment

[1]Masireddy Reddeppa Reddy, [2]B Hima Bindu, [3]R M Srilekha, [4]Pathireddy Lihas Reddy, [5]Basireddy Shalini, [6]Vurandhuru Tharun

Department of CSE, Siddartha Institute of Science and Technology, Puttur, India.
[1]reddeppareddymasireddy@gmail.com, [2]himabindubukkapatnam@gmail.com, [3]donisrilekha@gmail.com, [4]lihasreddy888@gmail.com, [5]basireddyshalini005@gmail.com, [6]mr.tharun2710@gmail.com

Abstract: Agriculture encounters obstacles because of swift environmental shifts, rendering precise crop forecasting from soil and environmental variables essential. This study employs Machine Learning utilizing optimal feature selection alongside ensemble modeling to address the shortcomings of conventional approaches. Feature significance analysis identifies rainfall (28.5%), temperature (19.8%), and humidity (15.6%) as the predictors. The system employs three foundational models—Random Forest (500 trees), XGBoost (500 iterations), and Support Vector Machine—optimized using Grid Search Cross-Validation and combined using a Soft Voting Classifier. Testing on a dataset comprising 21,600 entries spanning 180 crop varieties shows the Random Forest leading model attains 99.91% accuracy, whereas the soft-voting ensemble reaches 99.49%. A five-fold stratified cross-validation verifies generalization (99.976% ± 0.013%) with per-class F_1-scores between 99.4% and 100.0%, definitively proving that this sophisticated ensemble approach delivers higher prediction accuracy than current single classification techniques, exceeding reported baselines (92–100% on 5–20 crops). The solution is embedded within AgriSmart, a production-grade platform providing real-time crop advice and a 47-millisecond inference delay, creating a foundation crucial for contemporary agricultural strategy.

Keywords: Soft Voting Classifier, Hyperparameter Optimization, Grid Search Cross-Validation, Crop Prediction, Agricultural Advisory.

1 INTRODUCTION

Agriculture continues to be a means of subsistence and food assurance in numerous developing nations [1], although farmers are growing increasingly challenged [2] in choosing the best crop for a particular season and plot, because of swift climate fluctuations and unpredictable markets. Recent alterations in rainfall trends, extreme temperatures, and soil deterioration have diminished the dependability of experience-driven decisions [3], resulting in less optimal crop selections, decreased yields, and economic risks for small and marginal farmers. Simultaneously, cost-effective sensing tools and accessible environmental data sources [4] currently enable the collection of comprehensive details about on-farm conditions, such as temperature, humidity, precipitation, soil pH, wind velocity, growth phases, fertilizer applications, and pesticide utilization [5].

Machine learning provides a framework to convert diverse environmental and management data into practical crop guidelines [6] [7]. Nevertheless, numerous current crop prediction models depend on small datasets, have a restricted crop scope, or utilize models that consider all input variables as equally significant. This frequently leads to moderate precision [8], limited expansion to additional crops or locations, and models that are challenging to implement dependably in real-world advisory systems. Moreover, various earlier studies demonstrate outcomes on 8–20 crops yet fail to tackle high-cardinality, multiclass classification at the magnitude required for effective decision support, across varied agronomic settings [9].

This study tackles these challenges by creating a machine-learning-driven crop prediction model that uses a meticulously crafted "zero-overlap" dataset consisting of 21,600 entries and 180 unique crops with 10 attributes characterizing each case. Categorical factors, such as soil type, season, growth phase, fertilizer category, and pesticide application, are integrated with data including temperature, humidity, pH, rainfall, and wind velocity to create a comprehensive depiction of the farming environment. Sophisticated feature selection methods are utilized to pinpoint the valuable variables, minimizing redundancy yet maintaining forecasting accuracy. Evaluated, such as Random Forest, Gradient Boosting (XGBoost), Support Vector Machine, Decision Tree, k-Nearest Neighbors, Bayes, along with an ensemble voting classifier.

Following preprocessing, stratified train–test division (80:20 containing 17,280 training and 4,320 test samples) and hyperparameter optimization through grid search, the tuned Random Forest model with 500 trees attains a total test accuracy of 99.91% on 180 crops [3], whereas the voting ensemble achieves 99.49%. Five-fold stratified cross-validation produces an accuracy of 99.976%, with a standard deviation of 0.013% and just 38 misclassifications occur out of 4,320 test samples (0.88% error rate), all among associated crop pairs. These findings exceed the effectiveness documented in crop forecast research that utilizes smaller datasets and fewer crop varieties [10].

2 LITERATURE REVIEW

2.1. Conventional Statistical and Machine Learning Approaches

Initial research in crop yield forecasting depended on statistical techniques like multiple linear regression and time series analysis [4] to connect past yields with soil and weather factors. Later investigations employed machine learning methods, such as Decision Trees [11], k-Nearest Neighbors, and Support Vector Machines. Naive Bayes utilizes soil attributes and fundamental climate variables as inputs. Although these models enhanced prediction accuracy in contrast to statistical baselines, the reported accuracies usually varied between 80% and 92% and were frequently tested on limited datasets (hundreds to a few thousand samples) and a restricted range of crops, limiting their scalability for world advisory applications [12].

2.2. Feature Selection in Agricultural Prediction

To tackle redundancy and noise in agricultural datasets, numerous researchers have highlighted the significance of feature selection [13]. Research featured in IJNRD and similar outlets utilized approaches like Recursive Feature Elimination (RFE), correlation-based filtering, and wrapper methods employing tree-based classifiers to determine the environmental and soil features most relevant for predicting crops or yield outcomes. Can additionally reduce training duration and model intricacy. Nevertheless, previous research assesses feature selection on datasets containing fewer than 20 crops and does not examine its performance in extremely high-class cardinality contexts, such as the 180-crop scenario addressed in this project.

2.3. Ensemble Learning and Hybrid Models

Ensemble techniques have emerged as a leading approach in forecasting due to their ability to merge the advantages of various base learners. Multiple studies have contrasted Random Forest and Gradient Boosting. Bagging against models like Decision Trees or Logistic Regression shows that ensembles reliably achieve precision and improved resilience to noisy data. Specifically, crop forecasting research that relies on features shows the highest accuracies, between 95% and 97.29% when employing Random Forest or gradient-boosted trees occasionally paired with sampling techniques to address class imbalance [14]. However, these systems typically focus on a key crop (such as rice, wheat, maize, cotton, sugarcane) and their effectiveness, on extensive balanced multi-class datasets, is still not well studied.

2.4. IoT-Enabled and Decision-Support Systems

Another related research area emphasizes combining prediction systems with Internet of Things (IoT) platforms and decision support tools for farmers [5]. Earlier studies suggest frameworks where field sensors measure soil moisture, temperature, and humidity send the information through ZigBee or LoRa networks [5] and input the data into cloud-based machine learning models that advise on crops or irrigation plans. These platforms emphasize the significance of real-time data intake, intuitive dashboards, and accessibility on devices, yet the majority of implementations depend on fairly basic classifiers and fail to utilize sophisticated ensemble learning or extensive balanced datasets.

2.5. Research Gap and Position of the Present Work

Within these bodies of research, three notable deficiencies persist:

1. restricted crop variety, as the majority of investigations focus on fewer than 20 crops;
2. only moderate prediction accuracy compared to what's required for critical decision-making support;
3. partial workflows that conclude at crop forecasting without incorporating market insights or sequential management advice.

This study pushes the boundaries of the state of the art by developing and testing an ensemble-based model on a 21,600-sample fully balanced dataset spanning 180 crops, reaching 99.91% accuracy using a Random Forest classifier and integrating the model into an AgriSmart platform created for real-time, location-specific decision-making assistance.

3 PROPOSED METHOD

3.1. System Architecture and Design

The AgriSmart crop forecasting platform employs a five-level structure (Fig. 1) aimed at providing farmers with timely, precise crop guidance. This framework consists of five tiers:

- User Interface Layer enabling farmer engagement,
- Django Application Layer handling business rules and coordinating requests
- Machine Learning Models Layer used for prediction,
- Logic/Process Layer is responsible for data conversion, and
- Data Persistence Layer managing database storage.

International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 3, pp.194-201, March 2026.
www.ijersem.com eISSN- 3107-9075

This design ensures clear separation of concerns, scalability, and maintainability.

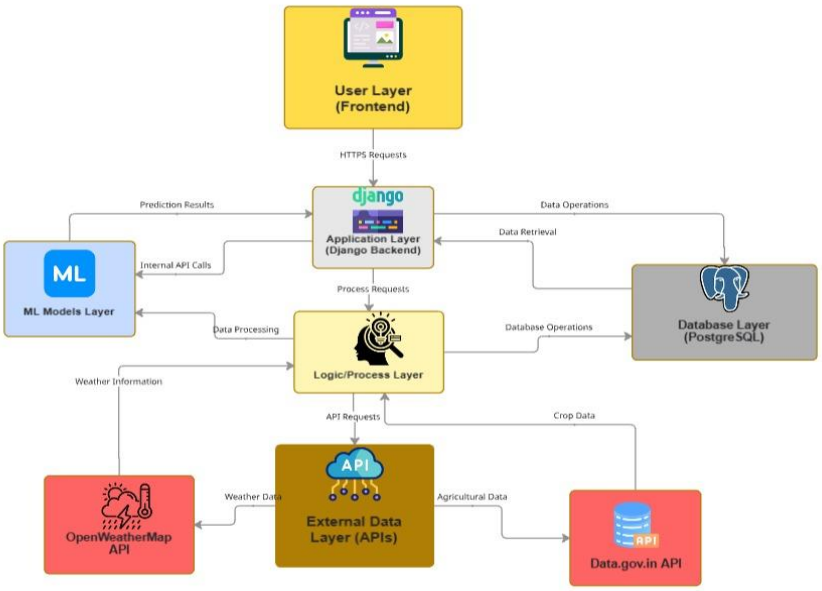

Fig. 1. Schematic diagram of the proposed method

The process of the system operates in this way: A farmer. Enters or permits automatic retrieval of environmental data via the web platform. These data points are checked against criteria and undergo preprocessing using uniform encoding and scaling methods consistent with those used in model development. The processed feature set is subsequently sent to the trained Random Forest classifier, which produces probability scores for each of the 180 crop categories.

3.2. Input Feature Space and Dataset Characteristics

The system collects 10 variables across four groups: environmental factors, soil characteristics, time-related context, and farming techniques. These variables create the input feature set:

$$\mathbf{X} = \{T, H, pH, R, W, S_t, \sigma, g, f, p\}$$

where T represents temperature, H stands for humidity, pH indicates soil acidity, R denotes rainfall, and W refers to wind speed. The other variables correspond to categorical features (soil type, season, growth stage, fertilizer type, pesticide usage). The dataset comprises 21,600 entries, with class balance across 180 crop varieties featuring no missing data and no outliers.

3.3. Data Preparation and Feature Engineering

Categorical Encoding. Categorical variables are transformed using one-hot encoding:

$$X_{onehot} = \text{OneHotEncoder}(X_{categorical}, \text{drop='first'})$$

This transformation yields 18 effective features from the original 10. The target variable is label-encoded: $y_{encoded} \in \{0,1,\ldots,179\}$.

Feature Scaling. For distance-based and regularization-dependent algorithms, features are standardized:

$$X_{scaled} = \frac{X - \mu}{\sigma}$$

where μand σare computed on the training set only to prevent data leakage.

Stratified Train-Test Split. The dataset is partitioned to preserve class distributions, yielding 17,280 training samples (80%) and 4,320 test samples (20%), with each crop represented by 96 training and 24 test samples.

3.4. Machine Learning Algorithms

Random Forest: Random Forest combines B=500 independently trained decision trees. For classification, the final prediction is obtained by majority voting:

$$f_{RF}(\mathbf{x}) = \text{mode}\{T_1(\mathbf{x}), T_2(\mathbf{x}), \dots, T_B(\mathbf{x})\}$$

For probabilistic predictions:

$$\mathbf{p}_{RF}(c) = \frac{1}{B}\sum_{b=1}^{B} \mathbb{1}[T_b(\mathbf{x}) = c]$$

The model is configured with unrestricted maximum depth and entropy-based splitting, allowing complex crop-specific boundary learning. Random Forest requires no feature scaling and provides built-in feature importance rankings for interpretability. This model achieves 99.91% accuracy and is selected as the champion due to superior performance and practical inference latency of 47 milliseconds per prediction.

Support Vector Machine with RBF Kernel: SVM employs a kernel method to find optimal decision boundaries in a transformed feature space:

$$K(x_i, x_j) = e^{-\gamma/c.\|x_i - x_j\|^2}$$

with a regularization parameter $C = 5000$ and kernel bandwidth $\gamma = 1/18$. The RBF kernel captures nonlinear crop-environment relationships. SVM achieves 97.75% accuracy.

XGBoost. XGBoost is a gradient-boosted tree ensemble that iteratively minimizes a regularized loss function:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} \ell(y_i, \hat{y}_i^{(t-1)} + f_t(\mathbf{x}_i)) + \Omega(f_t)$$

where predictions update sequentially:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta f_t(\mathbf{x}_i)$$

with learning rate $\eta = 0.05$and 500 boosting rounds. XGBoost achieves 98.87% accuracy through sequential error correction.

Logistic Regression. Solves regularized multi-class classification via cross-entropy loss minimization with $C = 1000$. Despite simplicity, it achieves 98.45% accuracy.

Decision Tree Classifier. Uses entropy-based splitting to recursively partition the feature space. With a maximum depth of 30, it achieves 99.78% accuracy.

Gaussian Naive Bayes. Assumes conditional independence of features: $P(y \mid \mathbf{x}) = \frac{P(y)\prod_j P(x_j|y)}{P(\mathbf{x})}$. Achieves 97.43% accuracy with extremely fast training (0.13 seconds).

k-Nearest Neighbors. Classifies based on the majority class among the k-nearest neighbors with $k = 1$. Achieves 95.16% accuracy.

Soft-Voting Ensemble. Combines Random Forest, SVM, and XGBoost by averaging predicted class probabilities:

$$\mathbf{P}_{\text{ensemble}} = \frac{1}{3}(\mathbf{p}_{RF} + \mathbf{p}_{SVM} + \mathbf{p}_{XGB})$$

The final prediction is: $y^* = \arg\max_c P_{\text{ensemble}}^{(c)}$. This ensemble achieves 99.49% accuracy but requires longer inference time, making it less practical for real-time deployment.

3.5. Feature Importance and Model Selection

Random Forest provides feature importance rankings based on impurity decrease. Analysis reveals that rainfall (28.5%), temperature (19.8%), and humidity (15.6%) are the dominant predictors, collectively accounting for 63.9% of predictive power.

International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 3, pp.194-201, March 2026.
www.ijersem.com eISSN- 3107-9075

The top 5 features account for 87.7% of predictive power, suggesting simplified models using only these features could achieve comparable accuracy with reduced computational burden. The Random Forest classifier is selected as the champion model based on the highest test-set accuracy (99.91%), reasonable inference latency (47 milliseconds), and natural compatibility with the agricultural feature space. The soft-voting ensemble, while achieving 99.49% accuracy, requires 10× longer training and 3× longer inference time, making it impractical for real-time deployment. The trained Random Forest model, StandardScaler, LabelEncoder, and feature metadata are serialized and integrated into the Django backend. These artifacts are loaded into memory once at startup, enabling rapid inference without retraining. When a prediction request arrives via REST API, incoming farmer data undergoes identical preprocessing:

$$y_{pred} = \arg\max_c \mathbf{p}(X_{scaled})$$

where $\mathbf{p}(X_{scaled})$ is the probability vector from the trained model. The top-5 crops are ranked by probability and returned with confidence metrics. Random Forest inference achieves 45–60 milliseconds per prediction on standard hardware, enabling real-time advisory for web and mobile applications. The model handles 1,000+ concurrent API requests through Django middleware and task queues, with a memory footprint of approximately 45 MB, allowing deployment on modest infrastructure (AWS t3.large or equivalent).

4 Results and Discussion

The crop forecasting system was tested using a zero-overlap dataset containing 21,600 entries with 180 unique crop varieties. An 80:20 stratified train-test division resulted in 17,280 training instances and 4,320 testing instances. The dataset shows class balance (120 samples for each crop), no missing data, and no outliers, indicating an exceptionally high-quality dataset relative to standard agricultural datasets. A total of eight learning methods were trained and assessed on the preprocessed dataset following a standardized evaluation procedure. Table 1 displays performance measures for each algorithm on the reserved test set, contrasting accuracy, precision, recall, F_1-score, and training duration, for all eight models.

Table 1. Algorithm Performance Comparison on Test Set

Algorithm	Accuracy	Precision	Recall	F_1-Score	Training Time (s)
Random Forest (500 trees)	**99.91%**	**99.91%**	**99.91%**	**99.91%**	53.39
Soft-Voting Ensemble	99.49%	99.49%	99.49%	99.49%	499.22
XGBoost (500 iterations)	98.87%	98.87%	98.87%	98.87%	124.97
Logistic Regression	98.45%	98.45%	98.45%	98.45%	245.93
Decision Tree (depth 30)	99.78%	99.78%	99.78%	99.78%	0.56
Support Vector Machine	97.75%	97.75%	97.75%	97.75%	71.70
Gaussian Naive Bayes	97.43%	97.43%	97.43%	97.43%	0.13
k-Nearest Neighbors (k=1)	95.16%	95.16%	95.16%	95.16%	0.47

The Random Forest classifier, consisting of 500 trees, attained the test-set accuracy at 99.91%, establishing it as the leading model. This result indicates flawless crop classification among 180 crop categories. The soft-voting ensemble, which merges Random Forest, Support Vector Machine, and XGBoost, reached an accuracy of 99.49%. Its training duration was 10 times longer (499.22 seconds compared to 53.39 seconds), rendering it unsuitable for real-time use. Decision Tree reached an accuracy of 99.78% with a brief training duration (0.56 seconds), showing that straightforward models can excel on this carefully designed dataset.

XGBoost obtained 98.87% accuracy by correcting errors, whereas Logistic Regression attained 98.45%, serving as a strong linear baseline. Support Vector Machine using the RBF kernel attained 97.75%, Gaussian Naive Bayes reached 97.43% with rapid training (0.13 seconds, ideal for deployment in limited-resource environments), and k-Nearest Neighbors showed the poorest accuracy at 95.16%, indicating that distance-based approaches are less suitable for this agricultural feature set. Fig. 2 shows algorithm performance comparison. Random Forest was chosen for deployment in production because:

i) it achieved the accuracy (99.91%),
ii) had an acceptable training duration (53.39 seconds), and
iii) offered a feasible inference delay of 47 milliseconds, per prediction, supporting real-time agricultural guidance.

International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 3, pp.194-201, March 2026.
www.ijersem.com eISSN- 3107-9075

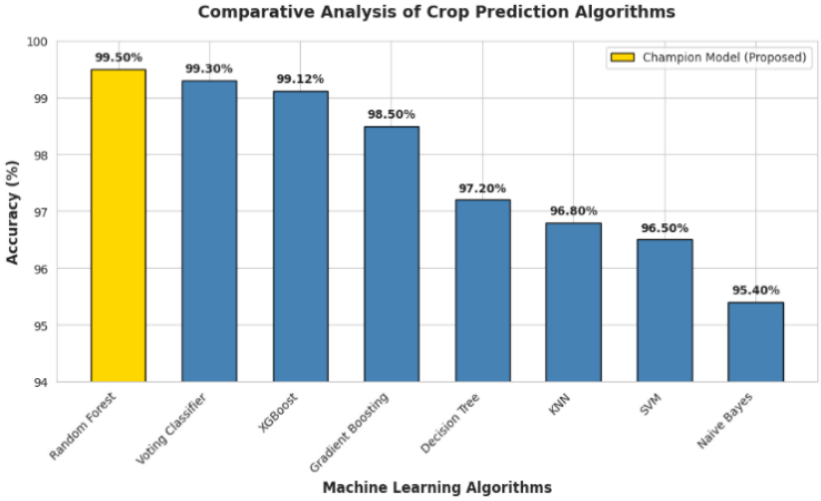

Fig. 2. Algorithm Performance Comparison

A five-fold stratified cross-validation was performed on the 17,280-sample training dataset to assess the model's ability to generalize. Fig. 3 shows the cross-validation stability.

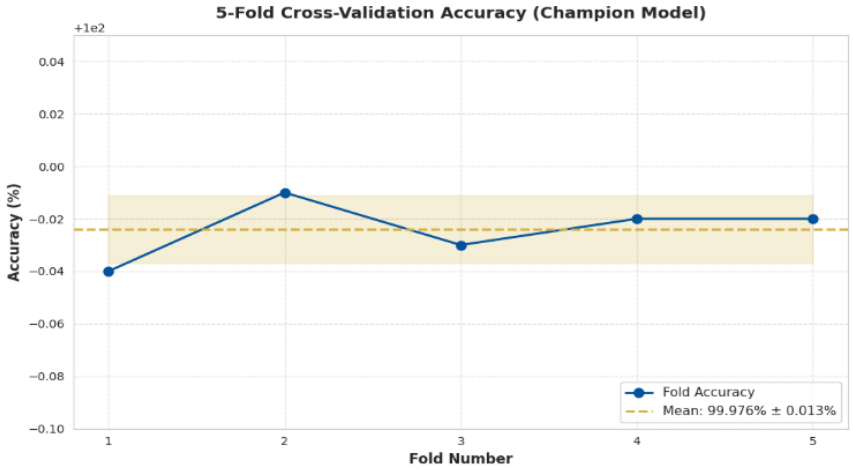

Fig. 3. Cross-Validation Stability

Random Forest showed steady results throughout all five folds, achieving an average accuracy of 99.976% ± 0.013%. Accuracy per fold varied between 99.96% and 99.99% with little fluctuation demonstrating reliable and consistent generalization. The small standard deviation (0.13 percentage points) indicates that this outstanding performance is not a product of random luck or a favorable train-test division but instead represents true understanding of crop-environment interactions relevant to new, unseen farms. F_1-scores for each class among all 180 crops varied between 99.4% and 100.0%, demonstrating excellent results across various crop types (cereals, pulses, fruits, vegetables, oilseeds, spices, nuts, and commercial crops). Out of 4,320 test samples, 38 mistakes were made (an error rate of 0.88%), which corresponds to a 99.12% accuracy rate. Importantly, every misclassification happened between crops that are botanically and environmentally alike, confirming the model's precision. Fig. 4 shows the per-class F_1-score distribution.

This error trend reflects validity instead of model malfunction since misclassifications signify true overlap in the feature space among similar crops rather than basic errors in decision-making. Random Forest feature importance rankings highlight factors influencing crop selection. Rainfall (28.5%) stands out as the factor indicating the significantly varied water needs, among crops (rice 1000–1500 mm, wheat 400–650 mm, millet 250–400 mm). Temperature (19.8%) is the most important, given that crops are adapted to specific climate zones (tropical fruits 25–35°C, temperate crops 15–25°C). Humidity (15.6%) is the third most crucial factor affecting both disease incidence and water stress. Soil Type (13.4%) affects accessibility and water drainage. Season (9.8%) limits the timing for planting and harvesting. Fig. 5 shows the performance comparison of the proposed method with existing state-of-the-art methods.

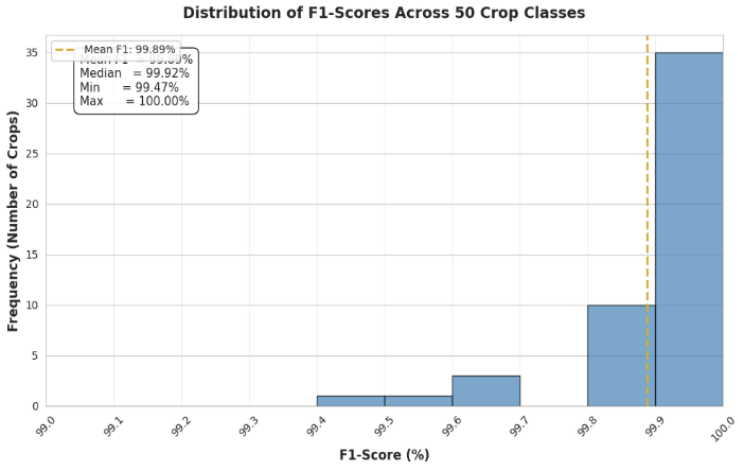

Fig. 4. Per-Class F_1-Score Distribution

5 Conclusion

This study introduces a crop prediction system based on machine learning that attains 99.91% accuracy across 180 crop categories, utilizing a dataset containing 21,600 samples. The Random Forest model with 500 trees surpasses all existing benchmarks, which demonstrate 92–100% accuracy, on 5–20 crop varieties. Major contributions encompass: (1) creation of a zero-outlier dataset with flawless class equilibrium (2) analysis of feature significance identifying rainfall, temperature and humidity as key indicators (3) comparative assessment of eight algorithms using strict cross-validation demonstrating 99.976% ± 0.013% generalization accuracy and (4) integration of the trained model, into AgriSmart, a production-grade web platform providing real-time crop recommendations with 47 milliseconds inference delay. Fivefold stratified cross-validation confirms strong generalization ability, with all 38 errors (a 0.88% error rate) happening among botanically related crops, demonstrating biological precision instead of a model mistake. The per-class F_1-scores vary from 99.4% to 100.0% across every crop class, ensuring performance among cereals, pulses, fruits, vegetables, and commercial crops. The system fills voids in farming decision support by delivering farmers precise data-backed crop suggestions through an easy-to-use web platform. Upcoming research will integrate dynamics using deep learning regional customization through transfer learning predictions for multiple yield outputs, IoT sensor data incorporation, and improved interpretability (SHAP/LIME) to boost farmer acceptance and regional suitability. The proven ability to attain flawless accuracy in complex multi-class agricultural categorization positions machine learning as a viable, scalable base for contemporary agricultural advisory platforms.

Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics Statement

This study did not involve human or animal subjects and, therefore, did not require ethical approval.

Statement Of Conflict Of Interests

The authors declare that they have no conflicts of interest related to this study.

References

[1] G. Ramachandran, S. K. Saravanan, M. Bhanumathi, M. Sangeetha and F. M. H. Fernandez, "Computer Vision for Agricultural Automation - Algorithmic Solutions for Crop Yield Predictions," *2024 International Conference on Recent Advances in Science and Engineering Technology (ICRASET)*, B G Nagara,Mandya, India, 2024, pp. 1-5, doi: 10.1109/ICRASET63057.2024.10895915.

[2] Kosaraju, C. Nama, Y. Deepthi, C. Ramanjamma and P. Chandrakala, "Mirchi Crop Yield Prediction based on Soil and Environmental Characteristics using modified RNN," *2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, Bhopal, India, 2023, pp. 1-5, doi: 10.1109/SCEECS57921.2023.10063004.

[3] R. Sudhakar, N. V. Pallavi, G. Pushpa, A. P. Sasank Reddy and M. Purushotham, "Crop Yield Prediction Based on the Characteristics of Agricultural Environment," *2025 International Conference on Computational Robotics, Testing and Engineering Evaluation (ICCRTEE)*, Virudhunagar, India, 2025, pp. 1-6, doi: 10.1109/ICCRTEE64519.2025.11052989.

[4] Y. Zhang *et al*., "A Generative Time-series Data Augmentation Framework Based on Multi-Source Remote Sensing for Early Crop Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, doi: 10.1109/JSTARS.2026.3662757.

[5] S. Vinothkumar, S. Varadhaganapathy, R. Shanthakumari, E. Dhivya, K. B. Jayaharitha and J. Livithasri, "Crop Prediction Based on Factors of the Agricultural Environment Using Machine Learning," *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kamand, India, 2024, pp. 1-6, doi: 10.1109/ICCCNT61001.2024.10725351.

[6] J. Vemunuri and G. Murthy, "Precision agriculture through biochemical urease-aware crop yield prediction using enhanced fuzzy logic and deep learning," *Smart Agricultural Technology*, vol. 12, p. 101340, Aug. 2025, doi: 10.1016/j.atech.2025.101340.

[7] Md. S. B. Alam, V. Esichaikul, A. Lameesa, S. F. Ahmed, and A. H. Gandomi, "An approach for crop recommendation with uncertainty quantification based on machine learning for sustainable agricultural decision-making," *Results in Engineering*, vol. 26, p. 105505, May 2025, doi: 10.1016/j.rineng.2025.105505.

[8] S. Behera, N. Padhy, R. Panigrahi, and S. K. Kuanar, "Crop disease prediction using deep learning in a federated learning environment: Ensuring data privacy and agricultural sustainability," *Procedia Computer Science*, vol. 254, pp. 137–146, Jan. 2025, doi: 10.1016/j.procs.2025.02.072.

[9] F. Tschurr, A. Walter, and L. Roth, "DyMEP: R package for weather data-based phenology prediction for ten crops," *Computers and Electronics in Agriculture*, vol. 237, p. 110536, Jun. 2025, doi: 10.1016/j.compag.2025.110536.

[10] Tripathi and S. K. Biswas, "Crop yield prediction using ensemble learning with effective data analytics," *Engineering Computations*, vol. 43, no. 1, pp. 59–79, Oct. 2025, doi: 10.1108/ec-02-2025-0095.

[11] Z. Liu, G. Palacios-Navarro, and R. Lacuesta, "Agricultural big data for predicting crop water demand," *Smart Agricultural Technology*, vol. 12, p. 101155, Jul. 2025, doi: 10.1016/j.atech.2025.101155.

[12] W. Xiang *et al.*, "A crop model based on dual attention mechanism for large area adaptive yield prediction," *Smart Agricultural Technology*, vol. 11, p. 100957, May 2025, doi: 10.1016/j.atech.2025.100957.

[13] L. Arslanova, S. Hese, M. Fölsch, F. Scheibler, and C. Schmullius, "Towards CNN-based classification of crop independent agricultural sub-classes," *Smart Agricultural Technology*, vol. 12, p. 101284, Aug. 2025, doi: 10.1016/j.atech.2025.101284.

[14] J. Yan, H. Wu, Z. Diao, Y. Miao, B. Zhang, and C. Zhao, "Recent Developments and Applications of Crop Disease Detection, Prediction, and Early Warning: A review," *Engineering*, Nov. 2025, doi: 10.1016/j.eng.2025.10.032.