International Journal of Convergent Technologies and Future Systems (IJCTFS)

Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php

Email: maulikavyanshuijctfs@gmail.com | ISSN No: 3108-1312

Research Today, Future Tomorrow Volume 02 Issue-01 January-February 2026

Multi-Algorithm Comparative Framework for Crop Recommendation Using Classical Machine Learning Models

Prem Jitendra Patil

Intern, Maulikavyanshu Foundation, Thalner Tal- Shirpur, Dist.- Dhule Maharashtra

Abstract:

The study called "Multi-Algorithm Comparative Framework for Crop Recommendation Using Classical Machine Learning Models" looks at how different Artificial Intelligence tools can help farmers figure out the crops to plant based on the soil and the weather.We tested six models to see which one works best. These models are Logistic Regression, Naive Bayes, SVM, KNN, Decision Tree . Naive Bayes is most accurate model. Logistic Regression and Decision Trees are easy to understand. On the hand SVM and KNN are better at finding complicated patterns in the data.This study shows that farming technology needs to be precise, fast and simple.The study "Multi-Algorithm Comparative Framework for Crop Recommendation Using Classical Machine Learning Models" provides a foundation, for work. This future work will use hybrid models to make farming systems more reliable.

Keywords:

Precision Agriculture, Crop Recommendation, Machine Learning (ML), Predictive Analytics, Hyperparameter Tuning, Model Interpretability, Gaussian Naive Bayes, Soil Chemistry, Climatic Data, Decision Support Systems

Introduction:

Agriculture is the primary lifesaving activity; multiple challenges limit its continuation. They include climate change, soil degradation, the large population requiring food, and market unpredictability [10, 11]. Historically, crop selection was done selectively based on the farmer's experience, historical patterns, and the trial-and-error approach. Although these methods work, there are inadequate to support the farmer's crop selection in the prevailing dynamic and multifaceted environmental and soil factor scenarios. As such, data-based solutions are essential in assisting farmers to make appropriately informed and reliable crop selections [7, 8].

Machine learning, or ML, has emerged in recent years as a powerful technology to develop a more intelligent and adaptive agricultural sector [11]. ML algorithms can analyse large datasets encompassing crop performance, soil characteristics, and climatic data to expose patterns that are often beyond the experience of human observation. Based on these results, machine learning models can advise the best crops for a location which can help improve yield, minimize waste of resources, and enhance food security [7].

In this paper, five classical machine learning algorithms are used on a crop recommendation dataset to compare performances: Logistic Regression, Naive Bayes, Support Vector Machines (SVM), K-Nearest Neighbour (KNN), and Decision Tree. Each algorithm has its own trade-offs: Logistic Regression and Decision Trees contain interpretability, Naive Bayes will be the simplest, SVM and KNN can process complex data. This work will assess the performance of these algorithms under different parameterization, providing analysis of the trade-offs involved between accuracy vs efficiency vs interpretation. In addition, these results can help develop decision support systems in support of furthering precision agriculture and sustainable farming.

Objectives:

The key aim of this study is to analyse and evaluate six classical machine learning algorithms (Logistic Regression, Naive Bayes, Support Vector Machines (SVM), K-Nearest Neighbours)

(KNN), Decision Tree), and compare the algorithms for crop

recommendation. Specifically, the study aims to:

1. Assess the accuracy, efficiency, and interpretability of each algorithm: Evaluating how different models balance performance with the ability for farmers to understand the result [6].

2. Optimize the model's predictive ability by tuning the hyperparameters.

3. Evaluate the trade-offs between lightweight models and ensemble methods.

4. Provide insights into the development of feasible tech-based agricultural crop advisory

systems.

5. Easy to identify which crop is suitable for the land [5].

Literature Review :

Recent advancements in precision agriculture have shown that machine learning is a tool for optimizing crop recommendation systems. Traditional farming decision-making, which relies on trial and error and local experience is no longer enough due to increasing changes and soil degradation. As a result using different types of data such as soil nutrients, pH levels and weather patterns in predictive models has become crucial for ensuring food security.

Research has found that machine learning algorithms can spot patterns in data that people might miss. For example Sam and D'Abreo said that choosing the crop must consider both environmental factors and economic factors to be viable for farmers. A study in Information also noted that data-driven advisory systems are vital for using resources and maximizing crop yield.

The literature shows that various algorithms have been used, each with its pros and cons. Naive Bayes is often mentioned for being simple and fast while Decision Trees and Logistic Regression are preferred when its essential to understand how the model works. On the hand complex models like Support Vector Machines and K-Nearest Neighbors are recognized for their ability to map complex behaviors in soil health data. While ensemble methods like Random Forest often provide accurate results they are often criticized for being difficult to understand which can make it hard for policymakers and farmers to adopt them.

Despite these advancements a significant gap remains: many studies focus on algorithms without comparing them in a way that balances accuracy with explainability. Furthermore there is a lack of emphasis on tuning algorithms for regional datasets. This study addresses these limitations by comparing six algorithms evaluating their performance, efficiency and practical reliability for sustainable precision agriculture and machine learning.This study focuses on machine learning, precision agriculture and crop recommendation systems.The goal of this study is to improve machine learning, in precision agriculture.

Challenges:

1. Data Quality and Availability. We have a problem with datasets because they have missing values and noise and some classes do not have data. This makes our models not very reliable. We need data to make good models. Data Quality and Availability is an issue because we cannot make good models without good data.

2. Hyperparameter Tuning. Every algorithm has its settings that we need to choose. For example when we use Logistic Regression we need to choose the value of C. When we use KNN we need to choose the value of k. When we use SVM we need to choose the kernel type. Choosing these settings is very important for our model to work well. Hyperparameter Tuning is crucial for model performance.

3. Model Generalization. It is hard to train a model on one dataset and then use it in soil and climate and regional conditions. Farmers work in environments and we need our models to work well in all these environments. Model Generalization is a challenge because our models need to work in all environments.

4. Interpretability vs. Accuracy Trade-off. Some models like Naive Bayes are very accurate but hard to understand. Other models like Decision Trees and Logistic Regression are easier to understand. May not be as accurate. This makes it hard to work with farmers and policymakers because they need to understand our models. Interpretability vs. Accuracy Trade-off is an issue because we need to find a balance between accuracy and interpretability.

5. Computational Complexity. Some algorithms like SVM with RBF kernel or large KNN searches need a lot of computing power. This can be a problem in areas with resources like rural farming communities. Computational Complexity is a challenge because we need to use algorithms that do not require a lot of computing power.

6. Integration with Real-world Farming. For our model predictions to be useful to farmers we need to present them in a way that's easy to understand and use. Farmers and agricultural workers need to know how to use our models. Integration with Real-world Farming is very important because our models need to be useful to farmers.

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | ISSN No: 3108-1312
Research Today, Future Tomorrow Volume 02 Issue-01 January-February 2026

7. Environmental Variability. Farming is affected by changing factors, like weather and soil degradation and pests. These factors can be hard to capture in our datasets, which can make our model predictions less accurate. Environmental Variability is an issue because we need to account for these changes. We need to find a way to capture these changing factors in our datasets.

Methodology:

Table 1: KNN PREDICTION TABLE

KNN: (K-Nearest Neighbours) is a supervised learning algorithm that classifies data based on the majority class of its nearest K neighbours using distance measures.

No of Features	K Value	Train	Test	Accuracy (in %)
6	3	0.6	0.4	97.05 %
6	3	0.8	0.2	97.04 %
6	5	0.6	0.4	97.38 %
6	5	0.8	0.2	97.04 %
6	7	0.6	0.4	96.81 %
6	7	0.8	0.2	97.04 %

This table showing the model accuracy for different settings of K value, train-test split, and number of features,

Accuracy ranges 96.8% – 97.5%.

Best accuracy: 97.5% (K=3, Train/Test=0.6/0.4, 7 features).

Table 2: SVC PREDICTION TABLE

SVM: Support Vector Machine (SVM) is a supervised learning algorithm that finds the best boundary (hyperplane) to separate classes by maximizing the margin between data points.

No. of Features	Kernel	Train	Test	Accuracy (in %)
7	linear	0.8	0.2	97.95 %
7	rbf	0.8	0.2	96.14 %
7	poly	0.8	0.2	98.18 %
7	linear	0.6	0.4	97.84 %
7	rbf	0.6	0.4	96.82 %
7	poly	0.6	0.4	97.39 %
7	linear	0.7	0.3	97.88 %
7	rbf	0.7	0.3	96.36 %

7	poly	0.7	0.3	97.58 %

This table shows the Support Vector Machine (SVM) model performance with different kernel types (linear, polynomial, RBF), train-test splits, and 8 selected features.

Accuracy ranges 96.1% and 98.18%

Best accuracy: 98.18% (RBF kernel, train/test=0.8/0.2, 7 features).

Table 3: NAIVE BAYES PREDICTION TABLE

Naive Bayes: Naive Bayes is a classification algorithm based on Bayes' theorem that assumes feature independence and predicts the class with the highest probability.

No. of Features	Naive Bayes	Accuracy %
7	Gaussian NB	99.55 %
7	Bernoulli NB	44.24 %
7	Complement NB	33.33 %

This table shows the Naïve Bayes model performance with three variants (Gaussian NB, Bernoulli NB, Complement NB) using 7 features.

Best accuracy: 99.55% (Gaussian NB)

Other model (Bernoulli NB, Complement) achieved about 69% accuracy.

Table 4: DECISION TREE PREDICTION TABLE

Decision Tree: Decision Tree is a machine learning algorithm that splits data into branches based on feature decisions to predict outcomes. It creates a tree where each node represents a condition and leaves represent final predictions.

No. Of Features	Tune Parameters	Max_Depth	Accuracy (in %)
7	gini	Nan	98.64 %
7	gini	5.0	40.00 %
7	gini	10.0	97.88 %
7	gini	15.0	98.64 %
7	entropy	Nan	9818 %
7	entropy	5.0	86.82 %
7	entropy	10.0	98.33 %
7	entropy	15.0	98.18 %

This table shows the Decision Tree model performance with different splitting criteria (gini, entropy) and max_depth settings.

Accuracy range86.82% – 98.64%

Best accuracy 98.64% (criterion=gini, max_depth=NaN,15.0 (7 features))

Table 4: Logistic Regression Prediction Table

No. of Features	Tune Parameters	Solver	Accuracy
7	12	lbfgs	95.45 %
7	12	saga	94.39 %

Best Accuracy: 95.45 % (Tune Parameter: 12. Solver: lbfgs).

Result:

This study compared KNN, SVM, Naïve Bayes, Decision Tree and Logistic Regression as traditional machine learning methods for recommending crops based on seven characteristics: nitrogen, phosphorus, potassium, temperature, humidity, pH and rain fall. The Naïve Bayes technique performed best in accuracy with an outcome of 99.55%. The Decision Tree and SVM closely followed with 98.64 and 98.18% respectively. KNN approached an outcome of 97%. Logistic Regression was lower, with an accuracy of 95.45%. Therefore, the results indicate that Naïve Bayes and tree/kernel-based methods are the best at predicting appropriate crops which can be a solid base for precision agriculture**.**

Sr. No.	Model	Parameters	Train	Test	Accuracy (in %)
1.	KNN	K = 3, Features = 7	0.6	0.4	97.5 %
2.	SVM	RBF Kernel, Features =7	0.8	0.2	98.18 %
3.	Naive Bayes	Bernoulli NB	-	-	99.55 %
4.	Decision Tree	Criterion = gini, Max_Depth=NaN	-	-	98.64 %
5.	Logistic Regression	Tune Parameter=12	-	-	95.45 %

References:

1. Sam, S., & D'Abreo, S. M. (2025). Crop recommendation with machine learning: Leveraging environmental and economic factors for optimal crop selection. arXiv preprint arXiv: https://arxiv.org/abs/2505.21201

2. MDPI. (2025). Evaluation of Learning-Based Models for Crop Recommendation in Smart Agriculture. Information, 16(8), 632. https://www.mdpi.com/2078-2489/16/8/632

3. Patil, S., & Kumar, R. (2025). Smart Crop Selection: A Machine Learning-Based Decision Support System for Farmers. International Journal for Research in Applied Science and Engineering Technology https://www.researchgate.net/publication/393167343

4. Maazallahi, A., et al. (2024). Naïve Bayes and Random Forest for Crop Yield Prediction. arXiv preprint arXiv:2404.15392. https://arxiv.org/abs/2404.15392

5. Abid, A., & Ahmad, S. (2025). Incorporating soil information with machine learning for crop recommendation to improve agricultural output. Scientific Reports (Nature Portfolio). https://pmc.ncbi.nlm.nih.gov/articles/PMC11903965/

6. Joshi, A., & Sharma, P. (2025). Performance Metrics and Accuracy Assessment of Machine Learning Models for Crop Recommendation. IEEE Xplore. https://ieeexplore.ieee.org/document/10845992/

7. Reddy, V., & Rao, K. (2025). Evaluating Machine Learning Models for Crop Recommendation with Optimal Nutrient Deviation Features. IEEE Xplore. https://ieeexplore.ieee.org/document/10961764/

8. Singh, H., & Kaur, M. (2025). Enhancing precision agriculture through cloud-based transformative crop recommendation model. Scientific Reports. https://pmc.ncbi.nlm.nih.gov/articles/PMC11914076/

9. Nayak, R., et al. (2024). AI-based Smart Crop Recommendation System for Sustainable Agricultural Production: A Data-driven Approach to Minimize Resource Use and Maximize Yield. Journal of Agricultural Science (MASU). https://masujournal.org/view_journal.php?id=566

10. Gomez, L., & Miller, J. (2025). Crop Recommendation System Using Machine Learning and Hybrid CNN-LSTM Architecture. SciTePress Digital Library. https://www.scitepress.org/Papers/2025/138908/138908.pdf

11. Dhingra, S., et al. (2024). Crop Recommendation System using Machine Learning with Nitrogen, Phosphorus, and Potassium soil data. GRENZE International Journal of Engineering and Technology. https://thegrenze.com/pages/servej.php?id=3211