Science World Journal Vol. 19(No 4) 2024 https://dx.doi.org/10.4314/swj.v19i4.32
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

ADVANCED BAGGING ENSEMBLE TECHNIQUE FOR MULTI-CROP PREDICTIVE MODELING TO ENHANCE AGRICULTURAL DECISION-MAKING

*[1]Muhammad Umar Abdullahi, [2]Gilbert I.O. Aimufua, [2]Morufu Olalere, [2]Rashidah Funke Olanrewaju, [2]Binyamin Adeniyi Ajayi

[1]Department of Computer Science, Federal University of Technology, Owerri, Nigeria
[2]Department of Computer Science, Nasarawa State University, Keffi, Nigeria

*Corresponding Author Email Address: umarfaruq54@gmail.com Phone: +2347067237166

ABSTRACT
Crop production is a cornerstone of agriculture, significantly influencing economies and farmers' livelihoods. However, fluctuating environmental conditions complicate the selection of suitable crops, requiring expertise in factors such as soil type, climate, humidity, rainfall, and temperature. Existing crop recommendation models primarily focus on a limited range of crops, such as rice, maize, and wheat, which restricts their utility across varied agricultural settings. Additionally, these models often exhibit inconsistent accuracy and high false-positive rates, undermining their reliability for practical use. To overcome these challenges, this study proposes a Bagging-based ensemble model that integrates seven machine-learning algorithms: Decision Tree, Support Vector Machine, Logistic Regression, Naive Bayes, Random Forest, K-Nearest Neighbor, and XGBoost. Leveraging a dataset enriched with diverse environmental and soil features—using soil type encoding and feature normalization—the model captures complex relationships that influence crop suitability. The ensemble model demonstrates an outstanding 99.9% accuracy, with macro-average precision, recall, and F1 scores of 99%, surpassing traditional models in performance. This advanced predictive tool offers a robust and versatile solution, enabling accurate and adaptable crop recommendations to support farmers and agricultural stakeholders in diverse environmental conditions.

Keywords: Crop Recommendation, Ensemble Learning, Multi-Crop Prediction, Bagging, Machine Learning.

INTRODUCTION

The farming sector serves as a crucial engine of revenue generation, especially in developing nations in which a significant segment of the population depends on growing crops for earnings and nourishment (Vishnoi & Goel, 2024). The projected rise in demand for food worldwide, coupled with the challenges posed by changing weather patterns regarding crop production, necessitates the adoption of more effective methods of farming (Nazir *et al.,* 2024).

Crop recommendation systems serve as essential instruments for enhancing worker efficiency by providing farmers with data-driven insights regarding optimal crop selection based on ecological and economic factors (Abdullahi *et al.,*2024; Bhola & Kumar, 2024). These systems analyze factors such as soil composition, climate, and rainfall patterns to provide tailored advice, helping to boost yields while conserving resources (Ayoola *et al.,* 2024). However, despite their potential, existing crop recommendation systems face significant challenges, particularly in the context of multi-crop prediction, where the task is to recommend the best combination of crops for a given area (Reddy *et al.,* 2024; Na & Na, 2024).

Conventional predictive models, such as Decision Trees and Support Vector Machines (SVM), often struggle to capture the intricate relationships between these variables (Rozenstein *et al.,* 2024). Furthermore, the multi-crop recommendation problem adds another layer of difficulty, as models must provide recommendations that account for the needs of multiple crops, which may have competing requirements (Ma, Ritsema & Wang, 2024). This results in sub-optimal predictions, reducing the potential for maximizing land productivity. Additionally, agricultural datasets are often noisy and imbalanced, which exacerbates the problem (Turchetta *et al.,* 2022). Models trained on incomplete or skewed data tend to perform poorly when applied to new or unseen conditions (Pecher, Srba & Bielikova, 2024).

One of the major limitations of conventional single-model approaches is their tendency to over-fit the data (Depaoli, Winter & Liu, 2024). In agricultural applications, over-fitting can occur when a model becomes too finely tuned to the training data, losing its ability to generalize to real-world situations (Xiong *et al.,* 2024). This leads to inaccurate crop recommendations, which can waste resources and diminish yields (Kumar et al., 2024). Furthermore, these models often struggle with data variability, failing to adapt to different regions or environmental conditions (Fisher & Koven, 2020). The result is inconsistent performance across different farming contexts, making it difficult for farmers to rely on such systems for actionable insights (Simelton & McCampbell, 2021).

To overcome these challenges, ensemble learning has gained distinction as a solution. Ensemble methods combine the predictions of multiple models to create a more accurate and stable output than any single model could achieve alone (Abdullahi *et al.,*2024; Sagana *et al.,* 2024). Bagging (Bootstrap Aggregating), in particular, has emerged as a promising technique for improving predictive performance in complex domains like agriculture (Mohammad *et al.,* 2024). Bagging works by generating multiple versions of a model, each trained on a different subset of the data, and then averaging their predictions (Özbayrak, Foster & Pyrcz, 2024). This process reduces variance and mitigates the risk of over-fitting, resulting in more stable and accurate predictions. Bagging's ability to handle noisy and diverse datasets makes it especially well-suited for the unpredictable nature of agricultural data (Sharma *et al.,* 2024).

In the context of multi-crop recommendation, a Bagging-based ensemble model holds significant promise. By aggregating the outputs of multiple models, Bagging can effectively manage the non-linear relationships and high variability in agricultural data, delivering more reliable and accurate crop recommendations (Savaş, 2024). This is particularly important in multi-crop settings,

Science World Journal Vol. 19(No 4) 2024
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University
https://dx.doi.org/10.4314/swj.v19i4.32

where the system must balance the varying needs of different crops grown together in the same environment (Rudinskienė *et al.,* 2024). Moreover, the ensemble approach helps reduce the chances of over-fitting, ensuring that the model remains robust when applied to real-world scenarios (Coulibaly, 2024). These advantages make Bagging a powerful tool for addressing the limitations of conventional crop recommendation systems, enabling more precise, stable, and generalization predictions (Gatou *et al.*,2024).

This paper focuses on developing a Bagging-based ensemble model for a multi-crop recommendation, aiming to enhance predictive accuracy and stability in the face of agricultural data complexity. The proposed model is designed to help farmers make informed decisions, ultimately improving productivity, resource management, and sustainability. The model will be evaluated on an agricultural dataset, and its performance will be compared with traditional machine learning models to demonstrate its superiority in real-world applications. Through this research, the study contributes to advancing the use of machine learning in agriculture, addressing the growing need for innovative solutions to optimize food production in a changing global landscape.

MATERIALS AND METHODS

This study employed a machine pipeline methodology that encompasses several stages including data collection, data pre-processing, data splitting, model training, and evaluation. This design is illustrated in Figure 1. The subsequent sections of the design will be discussed separately to understand the processes involved in achieving the objectives of this study.

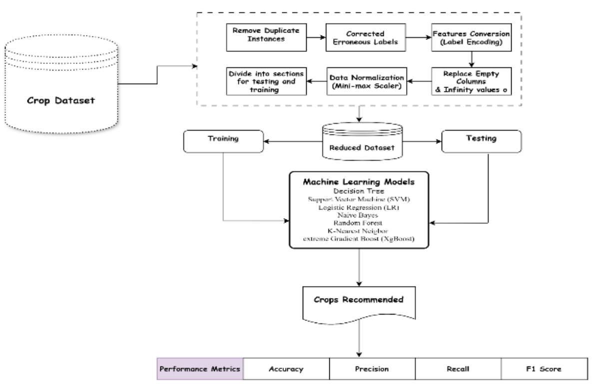

Figure 1**: Research Design**

Data Collection

The dataset used in this study was obtained from a secondary source on Kaggle. This dataset includes 2,200 instances with seven key features related to environmental and climatic conditions essential for crop growth. These features include temperature, rainfall, humidity, and soil composition elements like nitrogen, phosphorus, potassium, and pH. To enhance the dataset's relevance and specificity to the study objectives, new features were introduced, particularly soil types (such as clay, sandy, loamy, sandy-loam, loamy-sandy, etc.). These additional features were determined based on the unique requirements and properties of different crops, providing a more comprehensive dataset that aligns with the study's crop recommendation goals.

Table 1: Dataset Description

	Nitrogen	phosphorus	potassium	temperature	humidity	ph	rainfall	Clay
count	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000
mean	50.551818	53.362727	48.149091	25.616244	71.481779	6.469480	103.463655	0.500000
std	36.917334	32.985883	50.647931	5.063749	22.263812	0.773938	54.958389	0.500114
min	0.000000	5.000000	5.000000	8.825675	14.258040	3.504752	20.211267	0.000000
25%	21.000000	28.000000	20.000000	22.769375	60.261953	5.971693	64.551686	0.000000
50%	37.000000	51.000000	32.000000	25.598693	80.473146	6.425045	94.867624	0.500000
75%	84.250000	68.000000	49.000000	28.561654	89.948771	6.923643	124.267508	1.000000
max	140.000000	145.000000	205.000000	43.675493	99.981876	9.935091	298.560117	1.000000

Loamy	Silt	Sandy Loam	Loamy Sand	Silt loam	Clay loam	Drainage	WaterRetentionModerate	WaterRetentionHigh
2200.000000	2200.000000	2200.000000	2200.000000	2200.0	2200.000000	2200.000000	2200.000000	2200.000000
0.045455	0.863636	0.954545	0.909091	0.0	0.045455	1.863636	0.909091	0.409091
0.208346	0.343252	0.208346	0.287545	0.0	0.208346	0.343252	0.287545	0.491778
0.000000	0.000000	0.000000	0.000000	0.0	0.000000	1.000000	0.000000	0.000000
0.000000	1.000000	1.000000	1.000000	0.0	0.000000	2.000000	1.000000	0.000000
0.000000	1.000000	1.000000	1.000000	0.0	0.000000	2.000000	1.000000	0.000000
0.000000	1.000000	1.000000	1.000000	0.0	0.000000	2.000000	1.000000	1.000000
1.000000	1.000000	1.000000	1.000000	0.0	1.000000	2.000000	1.000000	1.000000

Data Preprocessing

In this study, data preprocessing played a pivotal role in transforming the raw dataset into a structured and ready-to-use format for training machine learning models. The process began by addressing common issues such as duplicate entries, which were removed to minimize noise and ensure cleaner data. Accuracy was a priority, so erroneous labels were corrected, and categorical variables like soil types were encoded into numerical values to ensure compatibility with the machine learning algorithms. Additionally, handling missing and infinity values was crucial for maintaining the integrity of the dataset—missing values were either imputed or discarded, while infinite values were set to zero, ensuring consistency throughout the dataset.

To further refine the data, data normalization was applied using a Min-Max scaler, which brought all feature values within a specified range of 0 to 1. This step was particularly beneficial for models sensitive to the magnitude of features, such as Support Vector Machine (SVM) and k-Nearest Neighbor (KNN). The normalization process ensured that all features contributed equally to the model's learning process, avoiding biases caused by differences in scale. Once preprocessing was complete, the dataset was split into training and testing sets to allow for unbiased model evaluation. Finally, a correlation matrix, depicted in Figure 2, was generated to examine the relationships between features, providing valuable insights into feature importance and interactions that could further enhance the model's performance.

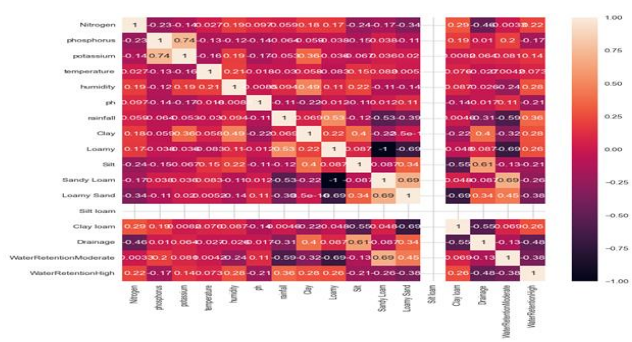

Figure 2: Correlation Matrix

Feature Extraction

In this study, feature selection was a critical step in narrowing down the most relevant environmental and soil-related factors for crop prediction. The primary focus was on variables like temperature, rainfall, humidity, and soil nutrient content, which includes key

Science World Journal Vol. 19(No 4) 2024 https://dx.doi.org/10.4314/swj.v19i4.32
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

elements such as nitrogen, phosphorus, potassium, and pH levels. These factors are known to significantly influence plant growth and crop yields. Additionally, a novel feature was introduced—soil types (e.g., clay, sandy, loamy)—which was derived from the specific properties of the crops. This new dimension of the dataset was essential for capturing the varying preferences of crops for different soil conditions, which could further enhance prediction accuracy.
To ensure the model was driven by the most impactful features, the study employed statistical techniques such as correlation analysis to evaluate the relationships between the selected features and the target variable (crop type). Features that showed a strong correlation to the target were retained, while those with minimal or no contribution were discarded, simplifying the dataset without sacrificing predictive power. The crop distribution across these selected features was then visualized using a boxplot (as shown in Figure 3), which provides a clear representation of how each factor varies across different crop groups. The plots revealed outliers in all factors, except nitrogen, underscoring the fact that plants have distinct optimal soil conditions. These outliers reflect the natural variation in how different crops respond to environmental and soil conditions, providing valuable insights for improving crop recommendations.

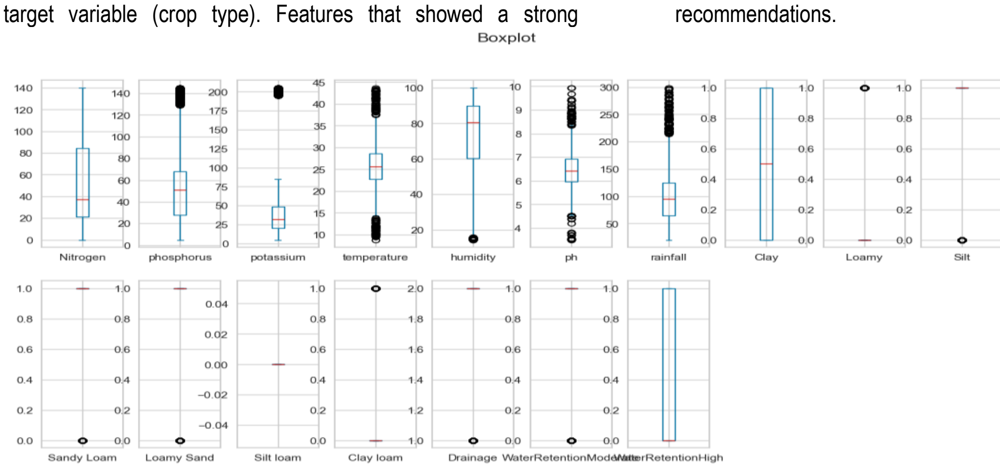

Figure 3: Outliers Detection of the Crop Distribution

Data Splitting
Data splitting is a fundamental step in machine learning workflows, ensuring that models are evaluated effectively on unseen data to assess their generalization capabilities. In this study, the dataset is divided into two main subsets: a training set, which comprises 80% of the total data, and a testing set, which accounts for the remaining 20%. This 80%-20% split ratio is a commonly adopted standard, striking a balance between providing enough data for the model to learn and leaving sufficient data for rigorous evaluation.
The training set is used to teach the models underlying patterns, relationships, and distributions within the data. By leveraging this subset, models iteratively adjust their parameters to minimize errors and improve predictions. The test set, on the other hand, serves as an unbiased dataset to evaluate how well the trained model can generalize to new, unseen data. This separation ensures that the model's performance metrics reflect its true predictive power and not its ability to memorize the training data.
This approach prevents overfitting, where a model performs excellently on the training data but poorly on new data. Additionally, by keeping the test set isolated throughout the training process, this study ensures an objective and reliable evaluation of the models' predictive accuracy, precision, recall, and other performance metrics. The 80%-20% split provides a robust foundation for assessing the individual model's ability to make accurate predictions while maintaining generalizability to real-world scenarios.

Model Training
In this research work, seven machine learning models i.e., Decision Tree, Support Vector Machine (SVM), Logistic Regression, Naive Bayes Random Forest K-Nearest Neighbor, and extreme Gradient Boosting (XGBoost) as an ensemble is utilized for crop recommendation. There are strengths and weaknesses of every model, and these give rise to a stronger prediction system when combined. The ensemble then trains an independent model over subsets of the training data taught to associate environmental conditions (for example temperature and humidity) with soil nutrients. You may need to use an ensemble method that uses multiple algorithms in the background, which aims at improving predictive accuracy by capturing different views of data.

Bagging (Bootstrap Aggregating)
This research deploys an ensemble-based Bagging (Bootstrap Aggregating) approach to enhance the crop recommendation model and prediction accuracy. As illustrated in Figure 4, the process involves first training arrays of base learners (B1 to B7) using several models with random forest, support vector machine, logistic regression, decision tree, XGBoost, K-nearest neighbor, and naive Bayes across combined portions of the training data. The subsets are created using bootstrapping, i.e., random sampling of the data with replacement to ensure that each model uses a slightly different variation of the dataset for training. After the prediction from each model is obtained, an aggregation method (usually majority voting) combines the predictions for the final output. This aggregated prediction uses the potential of each model in mitigating which reduces the overfitting and variance hence improving the predictiveness and robustness of crop recommendation.

Science World Journal Vol. 19(No 4) 2024 https://dx.doi.org/10.4314/swj.v19i4.32
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

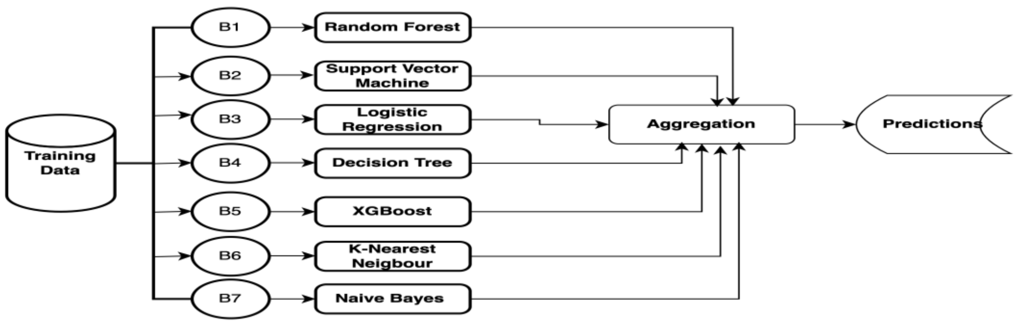

Figure 4: Bagging (Bootstrap Aggregating)

To indicate the models explicitly mathematically, we can label each base learner with its corresponding model.
Let the seven base models be denoted as:
M_1: Random Forest, M_2: Support Vector Machine (SVM), M_3: Logistic Regression, M_4: Decision Tree, M_5: XGBoost, M_6: K-Nearest Neighbor (KNN) and M_7: Naive Bayes.
Training Data: Let D represent the training dataset, where:

$$D = \{(x_i, y_i)\}_{i=1}^{N} \quad (1)$$

where x_i is the input feature vector, and y_i is the corresponding target value.
Base Learners: Each model M_j(where $j \in \{1,2,3,\ldots,7\}$) produces a prediction $\hat{y}_i^{(j)}$ for each input sample x_i:

$$\hat{y}_i^{(1)} = M_1(x_i),\ \hat{y}_i^{(2)} = M_2(x_i),\ \hat{y}_i^{(3)} = M_3(x_i),\ \hat{y}_i^{(4)} = M_4(x_i),\ \hat{y}_i^{(5)} = M_5(x_i),\ \hat{y}_i^{(6)} = M_6(x_i),\ \hat{y}_i^{(7)} = M_7(x_i) \quad (2)$$

Aggregation: After obtaining the individual predictions from all base learners, we aggregate these predictions to obtain the final prediction $\hat{y}_i^{(final)}$. The aggregation can be done as follows:
For classification (Majority Voting): $\hat{y}_i^{(final)} = mode(\hat{y}_i^{(1)}, \hat{y}_i^{(2)}, \ldots, \hat{y}_i^{(7)})$ (3)

For regression (Averaging):

$$\hat{y}_i^{(final)} = \frac{1}{7}\sum_{j=1}^{7} \hat{y}_i^{(j)} = \frac{1}{7}(\hat{y}_i^{(1)} + \hat{y}_i^{(2)} + \hat{y}_i^{(3)} + \hat{y}_i^{(4)} + \hat{y}_i^{(5)} + \hat{y}_i^{(6)} + \hat{y}_i^{(7)}) \quad (4)$$

Final Prediction: The final prediction for a given input sample x_i is obtained from the aggregated result as:

$$\hat{y}_i^{(final)} = f(\hat{y}_i^{(1)}, \hat{y}_i^{(2)}, \ldots, \hat{y}_i^{(7)}) \quad (5)$$

where f is the aggregation function (mode for classification, mean for regression).
Thus, the ensemble model can be described as:

$$\hat{y}_i^{(final)} = (M_1(x_i), M_2(x_i), \ldots, M_7(x_i)) \quad (6)$$

Crops Recommendation
The final recommendation provided by the Bagging-based ensemble model is which crops are best to be grown under specific Environmental and Soil conditions. We have already seen all these kinds of features (temperature, rainfall, humidity soil composition, and user-defined soil types) involved since the ensemble model is constructed by combining many machine learning models using a method called Bagging. The last prediction from this ensemble approach takes into account predictions outputted by all models so that the final recommendation is correct, stable, and robust. Moreover, this recommendation model suggests the crop type that has a greater likelihood of growing well with such environmental and soil features which will guide the farmer or agricultural planner in the decision-making. This variety of machine learning models in the ensemble creates a balanced recommendation by factoring different patterns within data that each model may overlook, and additionally, Bagging reduces overfitting/variance and increases the accuracy of these crop recommendations even more.

Performance Evaluation Metrics
Multiple criteria are frequently employed to evaluate the efficacy of a classification model. The metrics involve accuracy, precision, recall, and F1- score.

Accuracy
Accuracy is the percentage of correct predictions produced by the algorithm to the overall quantity of projections made.

$$\text{Accuracy (\%)} \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (7)$$

Precision
Precision is determined as the ratio of true positives to the total of true positives and false positives. The equation representing precision is as follows:

$$Precison = \frac{TP}{TP+FP} \quad (8)$$

Recall
The recall is a numerical statistic that determines the percentage of precisely detected positive occurrences that were incorrectly labeled as negative by the model. It is also known as the real positive rate. Mathematically, it is defined as the quotient derived by dividing the number of true positive (TP) occurrences by the total number of real positives and false negative (FN) cases.

Science World Journal Vol. 19(No 4) 2024
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University
https://dx.doi.org/10.4314/swj.v19i4.32

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

RESULTS AND DISCUSSION

Here, we provide the outcomes derived from the data assessment results. The findings are shown via the use of graphical representations and tabular data.

Classification Report of the Bagging Ensemble-Based Multi Crop Recommendation Model

The classification report for the crop recommendation ensemble-based prediction model is shown in Table 1. The classification report for the Bagging Ensemble-based multi-crop recommendation model showcases outstanding performance across various crops, evidenced by an overall accuracy of 99.9%. Each crop demonstrates high precision, recall, and F1-score metrics, with several crops like Rice, Kidneybeans, and Mungbean achieving perfect scores of 1.00 in all categories. While Chickpea shows a slight dip in recall (0.86), it still maintains a strong F1 score of 0.93. The macro and weighted averages, both at 0.99, further indicate the model's robust generalizability and reliability in predicting the suitable crops based on the given dataset, making it a highly effective tool for multi-crop recommendations.

Table 21: Classification Report of the Bagging Ensemble-Based Multi Crop Recommendation Model

Crop	Precision	Recall	F1-Score	Support
Rice	1.00	1.00	1.00	22
Maize	0.95	1.00	0.97	18
Chickpea	1.00	0.86	0.93	22
Kidney beans	1.00	1.00	1.00	15
Pigeon peas	0.95	1.00	0.97	18
Mothbeans	0.94	1.00	0.97	17
Mungbean	1.00	1.00	1.00	22
Black gram	1.00	1.00	1.00	29
Lentil	1.00	1.00	1.00	25
Pomegranate	1.00	1.00	1.00	20
Banana	1.00	1.00	1.00	18
Mango	1.00	1.00	1.00	20
Grapes	1.00	1.00	1.00	17
Watermelon	1.00	1.00	1.00	24
Muskmelon	1.00	1.00	1.00	24
Apple	1.00	1.00	1.00	26
Orange	1.00	1.00	1.00	15
Papaya	1.00	1.00	1.00	14
Coconut	1.00	1.00	1.00	19
Cotton	1.00	1.00	1.00	23
Jute	1.00	1.00	1.00	13
Coffee	1.00	1.00	1.00	19
Accuracy			99.9	440
Macro Avg	0.99	0.99	0.99	440
Weighted Avg	0.99	0.99	0.99	440

The table highlights the outstanding performance of the Bagging ensemble-based multi-crop recommendation model across a wide range of crops, including staples like rice, maize, chickpea, and lentil, as well as high-value crops such as mango, pomegranate, and coffee. With precision, recall, and F1-scores hovering around 1.0, the model demonstrates a remarkable ability to accurately predict crops suited for specific environmental and soil conditions. High precision across nearly all crops indicates a minimal occurrence of false positives, meaning the model is rarely incorrect in its recommendations. Similarly, strong recall values highlight the model's proficiency in identifying relevant crops, with minor exceptions, such as chickpeas, where a recall of 0.86 suggests a slight increase in false negatives.

Notably, the slight dip in F1-scores for crops like maize, pigeon peas, and moth beans (approximately 0.97) underscores the balance between precision and recall, indicating that while predictions are highly accurate, there may be room for fine-tuning. These slight variations suggest potential areas for improvement, such as enhanced feature selection or more granular data inputs to address the recall drop for chickpeas. Overall, the graph reflects a robust and reliable crop recommendation model that is poised to assist farmers and agricultural planners in making informed, data-driven decisions. Further refinements could elevate its already impressive performance, ensuring even greater applicability in diverse agricultural scenarios.

Confusion Matric of the Bagging Ensemble-Based Multi Crop Recommendation Model

The Bagging Ensemble-Based Multi Crop Recommendation Model for crop selection had excellent performance in all categories shown in Figure 5.

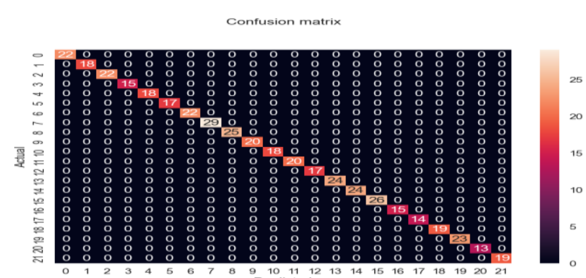

Figure 5: Confusion Matric of the Bagging Ensemble-Based Multi-Crop Recommendation Model

The confusion matrix presented in Figure 6 provides a compelling visual testament to the accuracy of the Bagging Ensemble-Based Multi-Crop Recommendation Model across 22 distinct crop categories. The dominance of high values along the diagonal reflects the model's ability to correctly classify a majority of crop types with exceptional precision. For instance, the model achieved perfect predictions for 22 instances of class 0 (rice), 18 instances of class 1 (maize), 29 instances of class 8 (blackgram), and 24 instances of class 15 (muskmelon). These results emphasize the model's effectiveness in distinguishing among diverse crop types based on environmental and soil conditions, reinforcing its robustness as a practical decision-support tool for agriculture.

While the confusion matrix reveals minimal misclassifications, such

https://dx.doi.org/10.4314/swj.v19i4.32

as 3 instances in class 5 and 4 instances in class 6 being assigned to incorrect categories, these errors are relatively rare compared to the overall performance. The model's ability to maintain such high accuracy across multiple crop classes demonstrates its adaptability and reliability, even in complex datasets with subtle variations. This precision in prediction not only aids farmers in selecting the best crops for specific conditions but also highlights the potential for scaling the model to broader agricultural contexts. With further optimization to address the few misclassifications, this model can become an indispensable resource for maximizing agricultural productivity and sustainability.

Learning Curve for Bagging Ensemble-Based Multi Crop Recommendation Model

The learning curve depicted in Figure 6 vividly captures the evolution of the Bagging Ensemble-Based Multi-Crop Recommendation Model's performance as the training set size increases. The red line, representing the training score, remains consistently high, hovering close to 1.0, demonstrating that the model achieves near-perfect accuracy on the training dataset. This constancy underscores the model's capacity to capture patterns within the training data effectively. On the other hand, the green line, which represents the cross-validation score, starts at a relatively lower value but climbs steadily as more training data is incorporated. This upward trend signifies that the model initially overfits, excelling on training data but struggling to generalize. However, with additional data, the cross-validation score approaches the training score, closing the performance gap.

This convergence of the two curves is a hallmark of improved generalization. It highlights the model's ability to perform well not only on the data it was trained on but also on unseen data, which is critical for real-world applications. The learning curve also suggests that the Bagging Ensemble approach benefits significantly from increased data, leveraging its ensemble nature to balance bias and variance effectively. This steady improvement in performance with more data reinforces the model's potential as a reliable tool for crop recommendation, capable of making accurate predictions across diverse environmental and soil conditions. The insights gleaned from this learning curve underscore the importance of robust training data to optimize the model's real-world applicability.

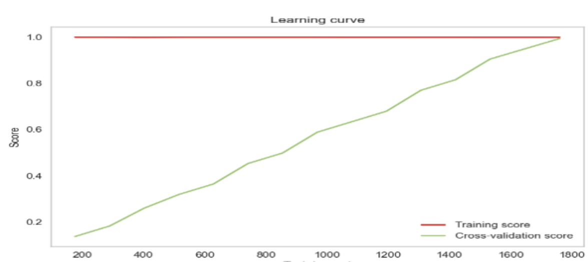

Figure 6: Learning Curve for Bagging Ensemble-Based Multi-Crop Recommendation Model

Residual Plot of the Bagging Ensemble-based Multi-Crop Recommendation Model

The residual plot shown in Figure 7 reveals a strikingly smooth and consistent pattern, illustrating the exceptional accuracy of the Bagging Ensemble-Based Multi-Crop Recommendation Model. The residuals, which represent the difference between the predicted and actual values, form a perfect horizontal line along the x-axis, with no noticeable deviation. This flatness signifies that the model has made precise predictions, with errors almost nonexistent. Such a residual plot indicates that the model has captured the underlying patterns in the data flawlessly, producing predictions that align almost perfectly with the actual outcomes.

This ideal residual pattern speaks volumes about the model's robustness and the quality of its predictions. A zero or near-zero residual across all predicted values suggests that the model has achieved a high level of fit to the data, with minimal to no bias or variance. In practical terms, it means that the Bagging Ensemble Model is highly reliable in crop recommendation, as it generates accurate predictions without significant errors. The residual plot thus provides compelling evidence of the model's effectiveness and its ability to make reliable crop predictions based on environmental and soil factors, reinforcing its potential for real-world agricultural applications.

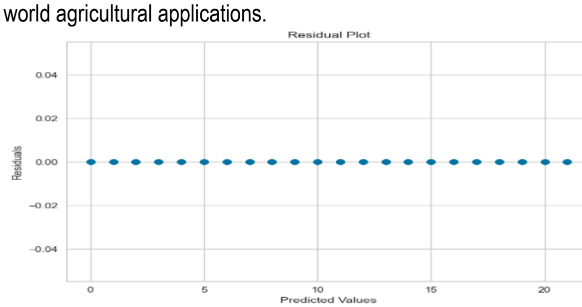

Figure 7: Residual Plot of the Bagging Ensemble-based Multi-Crop Recommendation Model

ROC Curve for the Bagging Ensemble-based Multi-Crop Recommendation Model

The ROC curve shown in Figure 8 highlights the Bagging Ensemble-Based Multi-Crop Recommendation Model's extraordinary ability to correctly classify crop types, boasting a perfect AUC score of 1.0 across all crop categories. This ideal score is a hallmark of flawless classification, where each crop class-specific curve reaches the upper-left corner of the plot, signifying a model that achieves a perfect balance between true positive rates and false positive rates. The model successfully distinguishes between crop types like Apple, Banana, Blackgram, Chickpea, and others, making precise recommendations that align closely with real-world agricultural needs. This perfect AUC value means that the model consistently classifies crop categories correctly, offering high reliability.

Further reinforcing the model's prowess is the macro-average AUC of 1.0, which indicates that, when aggregated across all classes, the model maintains impeccable performance. This perfect score signifies that no crop class was overlooked or misclassified, even when tested on out-of-sample data. The model's ability to generalize so effectively to unseen data speaks to its robust predictive capabilities, ensuring accurate crop recommendations in real-world conditions. This level of precision indicates that the model has a deep understanding of the complex relationships between environmental conditions and soil features, cementing its potential to guide agricultural decision-making with unparalleled accuracy.

Science World Journal Vol. 19(No 4) 2024
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University
https://dx.doi.org/10.4314/swj.v19i4.32

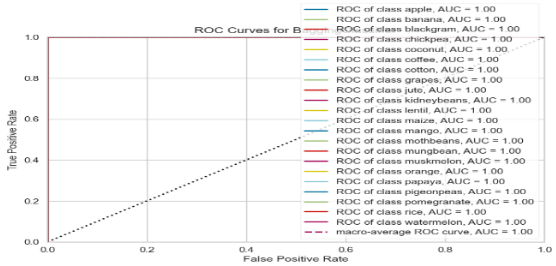

Figure 8: ROC curve for the Bagging Ensemble-based Multi-Crop Recommendation Model

Table 3: Performance of Selected Models

S/N	Model	Accuracy (%)	Precision	Recall	F Score	Support
1	KNN	98.4	0.98	0.99	0.99	141
2	DT	99.5	0.96	0.99	0.97	141
3	SVM	98.9	0.96	0.99	0.99	141
4	RF	99.8	0.99	0.99	0.99	141
5	LR	96.1	0.97	0.98	0.98	141
6	NB	99.6	0.99	0.99	0.99	141
7	XGBoost	89.1	0.89	0.88	0.88	141
8	Stacked Model	99.8	0.96	0.99	0.97	141
9	Bagging Model	99.9	0.99	0.99	0.99	141

As depicted in Table 2, each metric is expressed as a percentage to ensure consistency and ease of comparison. The models include KNN, Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), Naïve Bayes (NB), XGBoost, a Stacked Model, and a Bagging Model. The bars for each metric are color-coded: blue for Accuracy, orange for Precision, green for Recall, and red for F Score.
From the table, it is evident that most models exhibit high performance, with Accuracy and F Score values consistently exceeding 96% for the majority of models. The Bagging Model demonstrates the highest overall Accuracy at 99.9%, closely followed by Random Forest and the Stacked Model, both achieving 99.8%. Naïve Bayes also shows strong performance, with Accuracy nearing 99.6%. XGBoost, however, exhibits the lowest Accuracy at 89.1%, which is significantly below the other models, indicating potential challenges in its performance with this dataset.
Precision and Recall metrics are highly aligned for most models, with values consistently around 96% to 99%. Notably, Random Forest, Naïve Bayes, and the Bagging Model achieve the highest Precision and Recall at 99%, emphasizing their reliability in predicting both true positives and minimizing false negatives. XGBoost again trails, with Precision and Recall at 89% and 88%, respectively, reflecting reduced consistency in its predictions.
The F Score, a harmonic mean of Precision and Recall, reinforces the trends observed. Models such as Random Forest, Naïve Bayes, and the Bagging Model maintain F Scores of 99%, while XGBoost's F-Score drops to 88%. This disparity highlights XGBoost's relative underperformance compared to ensemble models like Random Forest and Bagging, which excel in balancing Precision and Recall.

Comparison with the Literature

The comparative analysis of this study against existing literature from Table 3 showcases the significant strides achieved in advancing machine learning-based crop recommendation systems.

DISCUSSION

Performance of Selected Models

Table 2 offers a detailed comparative analysis of each model's performance, providing a clear view of their strengths and suitability for crop prediction. Each predictive model was trained and evaluated individually to gauge its ability to forecast crop outcomes accurately.

Table 4: Comparison with the Literature

S/N	Author/Year	Machine Learning Model	Accuracy (%)
1	Prity *et al.* (2024)	LR, SVM, KNN, DT, RF, BG, AB, GB, and ET	14.1-99.31
2	Mahale *et al.* (2024)	LR, NB, SVM, KNN, DT, RF, BG, AB, GB, Bagging and Boosting	57.16-91.8
3	Subbulakshmi, Nirmaladevi & Rithani (2023).	XGBoost and MLP classifier algorithm	99.39
4	Nti *et al.* (2023)	AdaBoost GB, Light-GBM, RF, XGBoost, and Stacked TBEL	87.95–99.32
5	Our Study (2024)	KNN, DT, SVM, RF, LR, NB, XGBoost, Stacked Model, and Bagging Model	98.4-99.9

While previous studies, such as those by Prity *et al.* (2024) and Mahale *et al.* (2024), demonstrate a breadth of experimentation across multiple algorithms, their results highlight challenges with achieving consistently high accuracies across models. Prity *et al.*, for instance, report a wide accuracy range (14.1–99.31%), signifying variability in model performance likely influenced by dataset complexities and feature engineering. Similarly, Mahale et al.'s models achieved a maximum accuracy of 91.8%, underscoring limitations in generalization and model optimization.

 https://dx.doi.org/10.4314/swj.v19i4.32

In contrast, our study not only narrows the accuracy range (98.4–99.9%) but also demonstrates near-perfect precision, recall, and F-scores across the evaluated models. This success is attributed to a carefully curated preprocessing pipeline, the incorporation of Bagging and Stacked ensemble models, and the leveraging environmental and soil-related features that align directly with crop requirements. Subbulakshmi, Nirmaladevi, and Rithani's (2023) study, which achieved 99.39% accuracy using XGBoost and MLP classifiers, underscores the importance of sophisticated algorithms. However, our study surpasses their performance by achieving 99.9% with the Bagging model, demonstrating its robustness in handling diverse crop categories while mitigating overfitting.

A notable comparison is with Nti *et al.* (2023), who employed a stacked ensemble model (TBEL) and achieved up to 99.32% accuracy. While this showcases the power of ensemble techniques, our study further pushes the boundary by combining Bagging's stability with Stacked Model's ability to aggregate diverse predictions, leading to a higher peak accuracy. The combination of strong preprocessing, targeted feature selection, and cutting-edge ensemble learning positions our study as a benchmark for crop recommendation systems, proving its potential to outclass existing methodologies and set a new standard for machine learning applications in agriculture.

Conclusion

This study successfully developed a Bagging ensemble-based multi-crop recommendation model by combining seven machine learning algorithms—Decision Tree, Support Vector Machine, Logistic Regression, Naïve Bayes, Random Forest, K-Nearest Neighbor, and XGBoost—into a robust ensemble framework. The Bagging-based model achieved an outstanding accuracy of 99.9%, with precision, recall, and F-scores reaching 99%, signifying its capacity to provide accurate crop recommendations with minimal misclassification. The preprocessing pipeline, including one-hot encoding of soil types, normalization using the Min-Max scaler, and feature selection targeting critical factors such as nitrogen, phosphorus, potassium, pH, and soil types, significantly contributed to the model's superior performance.

A pivotal innovation lies in the integration of novel agronomic features such as Drainage: WaterRetentionModerate and WaterRetentionHigh, along with enriched soil classifications like Clay, Loamy, Silt, Sandy Loam, Loamy Sand, Silt Loam, and Clay Loam. These additions unlock a deeper understanding of soil and environmental dynamics, enhancing the model's ability to predict crop suitability with unparalleled precision. By reducing false positives and false negatives while significantly improving true positive rates, this model ensures dependable recommendations, empowering farmers with scientifically validated insights for optimized crop selection.

In comparison with the literature, this research stands out by addressing limitations in previous studies. While works like Prity et al. (2024) and Mahale et al. (2024) offered accuracies up to 99.31% and 91.8%, their methodologies lacked the innovative feature engineering and ensemble-driven framework presented here. Similarly, Subbulakshmi et al. (2023) and Nti et al. (2023) delivered high-performance models, yet their scope fell short of encompassing the diverse agronomic factors captured in this study. By surpassing these benchmarks, this research sets a new gold standard for crop recommendation systems.

This study does not just improve prediction accuracy; it establishes a transformative approach for decision-making in agriculture. The model's scalability and adaptability make it a powerful tool for addressing varying soil profiles, environmental conditions, and dynamic agronomic challenges. It provides a pathway for sustainable farming practices, empowering stakeholders to harness the potential of precision agriculture for improved crop productivity and food security.

Future Work

Looking forward, future expansions could include real-time deployment, dynamic factor integration, and testing across diverse geographies, further cementing the model's relevance and impact.

Conflict of Interest: The corresponding author, representing all the contributions, confirms the absence of any disputes of interest.

REFERENCES

Ayoola, A. J., Essien, J., Ogharandukun, M., & Uloko, F. (2024). Data-Driven Framework for Crop Categorization using Random Forest-Based Approach for Precision Farming Optimization. *European Journal of Computer Science and Information Technology*, *12*(3), 15-25.

Bhola, A., & Kumar, P. (2024). Farm-Level Smart Crop Recommendation Framework Using Machine Learning. *Annals of Data Science*, 1-24.

Coulibaly, A. (2024). *Smart Farming: Computer Simulation and Predictive Model for Cassava* (Doctoral dissertation, Université d'Ottawa/University of Ottawa).

Depaoli, S., Winter, S. D., & Liu, H. (2024). Under-Fitting and Over-Fitting: The Performance of Bayesian Model Selection and Fit Indices in SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, *31*(4), 604-625.

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2018MS001453.

Gatou, P., Tsiara, X., Spitalas, A., Sioutas, S., & Vonitsanos, G. (2024). Artificial Intelligence Techniques in Grapevine Research: A Comparative Study with an Extensive Review of Datasets, Diseases, and Techniques Evaluation. *Sensors*, *24*(19), 6211.

Jenkins, M., & Block, D. E. (2024). A Review of Methods for Data-Driven Irrigation in Modern Agricultural Systems. *Agronomy*, *14*(7), 1355.

Kumar, M., Agrawal, Y., Adamala, S., Pushpanjali, Subbarao, A. V. M., Singh, V. K., & Srivastava, A. (2024). Generalization Ability of Bagging and Boosting Type Deep Learning Models in Evapotranspiration Estimation. *Water*, *16*(16), 2233.

Ma, S., Ritsema, C. J., & Wang, S. (2024). Achieving sustainable crop management: A holistic approach to crop competitiveness assessment and structure optimization with dual natural-social environmental impacts. *Agricultural Systems*, *218*, 103968.

Mahale, Y., Khan, N., Kulkarni, K., Wagle, S. A., Pareek, P., Kotecha, K., ... & Sharma, A. (2024). Crop recommendation and forecasting system for Maharashtra using machine learning with LSTM: a novel expectation-maximization technique. *Discover Sustainability*, *5*(1), 134.

Mohammad, S. I., Vani, K. S., Lokeshwar, G., & Lakshmi, K. V.

(2024, March). Optimizing Crop Selection Using Bagging and Boosting Models Based on Soil Composition and Environmental Factors. In *2024 International Conference on Distributed Computing and Optimization Techniques (ICDCOT)* (pp. 1-8). IEEE.

Na, M. H., & Na, I. S. (2024). AI-Powered Predictive Modelling of Legume Crop Yields in a Changing Climate. *Legume Research: An International Journal*, *47*(8).

Nazir, M. J., Li, G., Nazir, M. M., Zulfiqar, F., Siddique, K. H., Iqbal, B., & Du, D. (2024). Harnessing soil carbon sequestration to address climate change challenges in agriculture. *Soil and Tillage Research*, *237*, 105959.

Özbayrak, F., Foster, J. T., & Pyrcz, M. J. (2024). Spatial bagging to integrate spatial correlation into ensemble machine learning. *Computers & Geosciences*, *186*, 105558.

Pecher, B., Srba, I., & Bielikova, M. (2024). A survey on the stability of learning with limited labelled data and its sensitivity to the effects of randomness. *ACM Computing Surveys*, *57*(1), 1-40.

Prity, F. S., Hasan, M. M., Saif, S. H., Hossain, M. M., Bhuiyan, S. H., Islam, M. A., & Lavlu, M. T. H. (2024). Enhancing Agricultural Productivity: A Machine Learning Approach to Crop Recommendations. *Human-Centric Intelligent Systems*, 1-14.

Reddy, G. V., Reddy, M. V. K., Spandana, K., Subbarayudu, Y., Albawi, A., Chandrashekar, R., ... & Praveen, P. (2024). Precision farming practices with data-driven analysis and machine learning-based crop and fertiliser recommendation system. In *E3S Web of Conferences* (Vol. 507, p. 01078). EDP Sciences.

Rozenstein, O., Cohen, Y., Alchanatis, V., Behrendt, K., Bonfil, D. J., Eshel, G., ... & Lowenberg-DeBoer, J. (2024). Data-driven agriculture and sustainable farming: friends or foes?. *Precision Agriculture*, *25*(1), 520-531.

Rudinskienė, A., Marcinkevičienė, A., Velička, R., & Steponavičienė, V. (2024). The effects of incorporating caraway into a multi-cropping farming system on the crops and the overall agroecosystem. *Agronomy*, *14*(3), 625.

Sagana, C., Devi, R. M., Sangeetha, M., Kumar, K. K., Lalit, K., & Sameera, N. M. (2024, March). Ensemble Model-Based Crop Recommendation System with Data Generation Using GAN. In *International Conference on Artificial Intelligence and Smart Energy* (pp. 310-321). Cham: Springer Nature Switzerland.

Savaş, S. (2024). Application of deep ensemble learning for palm disease detection in smart agriculture. *Heliyon*, *10*(17).

Sharma, M., Mittal, N., Mishra, A., & Gupta, A. (2024). Ensemble learning for prominent feature selection and electric power prediction in agriculture sector. *Multimedia Tools and Applications*, 1-28.

Simelton, E., & McCampbell, M. (2021). Do digital climate services for farmers encourage resilient farming practices? Pinpointing gaps through the responsible research and innovation framework. *Agriculture*, *11*(10), 953.

Subbulakshmi, B., Nirmaladevi, M., & Rithani, R. (2023). Crop Prediction and Recommendation Using Ensemble of DL Models. *Mapana Journal of Sciences*, *22*.

Turchetta, M., Corinzia, L., Sussex, S., Burton, A., Herrera, J., Athanasiadis, I., ... & Krause, A. (2022). Learning long-term crop management strategies with cyclesgym. *Advances in neural information processing systems*, *35*, 11396-11409.

Vishnoi, S., & Goel, R. K. (2024). Climate smart agriculture for sustainable productivity and healthy landscapes. *Environmental Science & Policy*, *151*, 103600.

Xiong, H., Li, J., Wang, T., Zhang, F., & Wang, Z. (2024). EResNet-SVM: an overfitting-relieved deep learning model for recognition of plant diseases and pests. *Journal of the Science of Food and Agriculture*, *104*(10), 6018-6034.