



ML-CSFR: A UNIFIED CROP SELECTION AND FERTILIZER RECOMMENDATION FRAMEWORK BASED ON MACHINE LEARNING

AMIT BHOLA* AND PRABHAT KUMAR†

Abstract. Sustainable and substantial crop production is essential globally, especially considering the increasing population. To achieve this, selecting appropriate crops and applying necessary fertilizers are pivotal for ensuring satisfactory crop growth and productivity. Farmers have relied heavily on intuition when choosing which crops to cultivate and suitable fertilizers to use in a given season. However, this traditional approach often needs to consider the significant impact of current environmental and soil conditions on crop growth and yield. Overlooking these factors can have far-reaching consequences, impacting not just individual farmers and their households but also the entire agricultural sector. The integration of machine learning offers a promising avenue for addressing these challenges and providing practical solutions. The core contribution of this research lies in proposing a unified framework termed Machine Learning-enabled Crop Selection and Fertilizer Recommendation (ML-CSFR). This framework's primary objective is to predict appropriate crops accurately and subsequently suggest corresponding fertilizers based on specific agricultural conditions. The initial phase involves the identification of proper crops for individual farmlands, considering local input variables. This phase employs artificial neural networks (ANN) to filter crops effectively using the available choices. The next phase utilizes soil and environmental parameters to anticipate the optimal fertilizer for the selected crops. This phase leverages the XGBoost (XGB) model to predict the most suitable fertilizers accurately. This two-phase approach ensures a comprehensive and effective recommendation system for enhancing agricultural outcomes. Experimental results demonstrate the effectiveness of this framework, achieving an accuracy score of 99.10% using ANN and 97.66% for XGB. The framework's capability to deliver tailored recommendations for individual farms and its potential to integrate real-time sensor data positions it as an effective tool for improving agricultural decision-making.

Key words: Machine learning, Fertilizer recommendation, Digital agriculture, Crop selection, Neural network

1. Introduction. Agriculture stands as the foundation of our society, fulfilling the nutritional requirements of billions worldwide. With the relentless growth of the world's population, the assurance of a consistent and reliable food source has become paramount. The significant role of agriculture is highlighted by its substantial contribution to India's GDP, amounting to 18.3% during the fiscal year 2022-2023 [1]. Additionally, this sector serves as the source of livelihood for approximately 50% of the country's workforce [2]. Despite its importance, the agricultural industry has experienced a decline in its performance in recent times. According to the Food and Agriculture Organization of the United Nations (FAO), approximately one-third of all food produced for human consumption across the globe is lost or wasted each year [3]. Insufficient land holdings, soil depletion, improper crop choices, inadequate fertilization, climate variations, and plant disease influence these losses.

Emerging technologies like Machine Learning (ML) [4], [5], Deep Learning (DL) [6], [7], and the Internet of Things (IoT) [8] have proven highly advantageous for the agricultural sector. These advancements offer the potential to enhance productivity and simultaneously ensure ecological sustainability by fortifying conventional farming practices. However, many small and marginal farmers still employ traditional or customary agricultural methods. For instance, these farmers often rely on their rudimentary knowledge to select crops and suitable fertilizers, usually favoring traditional or popular crops within their locality. Consequently, this reliance on traditional methods can compromise crop yields and soil fertility [11] due to insufficient scientific insights [15]. An adverse consequence of such practices is increased soil acidity, inappropriate selection and crop-specific fertilizers, and inadequate soil nutrient management. Additionally, the quality and productivity of crops are influenced by environmental conditions, climate variations, soil attributes, and water levels, further underlining the complexity of agricultural outcomes.

*Computer Science and Engineering Department, National Institute of Technology Patna, Bihar, India. (amitb.phd19.cs@nitp.ac.in).

†Computer Science and Engineering Department, National Institute of Technology Patna, Bihar, India. (prabhat@nitp.ac.in).

Selecting suitable crops and appropriate fertilizers tailored to filtered crops is pivotal in augmenting agricultural output and enhancing quality. Driven by the challenges above, this study aims to identify and tackle the intricacies of crop production by proposing a machine learning enabled Crop Selection and Fertilizer Recommendation (ML-CSFR) framework. This framework is designed to navigate the complex decision-making process of crop selection and appropriate fertilizer recommendation. It achieves this by considering various influential factors, including temperature, humidity, pH, rainfall, and soil nutrient levels. Notably, soil nutrients like Nitrogen (N), Phosphorus (P), and Potassium (K) are paramount for promoting plant growth and enhancing yield [16]. Concurrently, pH governs chemical reactions within the soil by determining its acidic or alkaline nature. Moreover, the development of plants is significantly influenced by electrical conductivity (EC), which also indicates soil fertility, water quality, and salinity.

The government has taken initiatives to enhance agricultural productivity by issuing Soil Health Cards (SHC) to individual farmers after assessing their farm's soil composition. These cards provide details about the soil's macro and micro nutrient levels. However, the conventional farming practices followed by farmers often fail to capitalize on this valuable information, resulting in suboptimal agricultural productivity. The ML-CSFR framework is proposed as an accessible and cost-effective solution to address these challenges. This framework suggests crops and corresponding fertilizers based on localized parameters by leveraging machine learning techniques. The ML-CSFR framework encompasses two distinctive phases: crop filtration and fertilizer recommendation. In the initial phase, crops are filtered based on local and regional variables. This filtration process involves the selection of 'n' crops that align with the soil and weather conditions while excluding those less suitable. Subsequently, tailored fertilizer recommendations are provided for each filtered crop in the second phase, guided by the soil's specific parameters. This research offers several noteworthy contributions:

1. The study introduces a two-phase Machine Learning-based Crop Selection and Fertilizer Recommendation (ML-CSFR) framework designed to provide farmers with optimal crop selection and fertilizer recommendations to enhance their returns. The initial phase of this framework involves assisting in selecting appropriate crops by filtering them based on the specific soil nutrient levels and the prevailing regional weather conditions associated with each farmland. Subsequently, the second phase generates fertilizer predictions for each filtered crop, considering the localized soil parameters.
2. Extensive experiments are conducted to demonstrate the efficacy of the proposed framework.
3. The first phase is evaluated on six benchmark models, including Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbour (KNN), XGBoost (XGB), and Artificial Neural Networks (ANN), for crop filtration. The evaluation results highlight that the proposed Artificial Neural Network (ANN) model attains the highest accuracy scores of 99.07% for validation and 99.10% for testing, while 99.13% precision and 99.24% recall, surpassing the performance of all other studied models.
4. The second phase evaluates eight benchmark models for fertilizer prediction: Logistic Regression (LR), DT, Gaussian Naive Bayes (GaNb), LrSVM, XGB, and RF. The evaluation results emphasize that the proposed XGB model excels, achieving outstanding accuracy scores of 99.23% for training and 97.66% for testing.
5. This empirical analysis validates that the proposed crop filtration in the first phase aligns with improved crop intensity and productivity. Furthermore, the framework's fertilizer recommendations in the second phase undergo validation through credible sources referenced in citations, thereby justifying the credibility and reliability of the framework's output.

The subsequent sections of this paper are organized as follows: In Section 2, a comprehensive overview of related works is provided. Section 3 outlines the methodology for the proposed framework. Section 4 discusses the experimental setup, data analysis, and methods used. The results and discussion are presented in Section 5. Lastly, Section 6 concludes the findings and discusses potential avenues for future research.

2. Related works. Numerous researchers have made contributions to devising diverse solutions to enhance crop productivity. Machine learning [10] and Deep learning have made their way into agriculture to enable more innovative and accurate crop selection decisions and improve yield through precise fertilizer recommendations [12]. Automated systems equipped with IoT [27] are also facilitating farmers in monitoring processes and receiving alerts, including weather conditions and soil moisture updates [17].

Agriculture has seen notable advancements in its pursuit of real-time crop selection [36] and fertilizer prediction [13], [14]. Cheema et al. [18] introduced a diverse crop model leveraging multiple soil parameters for

identifying suitable crops. The proposed approach harnessed a Quantum Value-based Gravitational Search Algorithm (GSA) to discover optimal solutions. Soil attributes such as pH, salinity, texture, nitrogen, phosphorous, and potassium were explored and employed as inputs for crop selection. Bakhavatchalam et al. [19] presented a crop prediction system that relies on various attributes and utilizes a multilayer perceptron (MLP), JRip, and decision table classifier. A range of machine learning models was implemented on the WEKA platform, with the most effective MLP model achieving an accuracy of 98.22%. Jain et al. [20] devised a soil-based machine learning comparative analytical framework for predicting crop yield production. This framework employed soil characteristics and climatic factors to categorize crop yield predictions as high, low, or medium. Gupta et al. [21] introduced a crop recommendation system integrating MapReduce and K-means clustering. This model considered crop yields per acre for distinct regions based on different varieties cultivated in the target zone. Mariammal et al. [22] proposed a feature selection method called Modified Recursive Feature Elimination (MRFE) for crop prediction, aiming to select vital features from crop data. This method employed a ranking algorithm to identify significant attributes. The results highlighted that MRFE outperformed various wrapper-based feature selection techniques, achieving an accuracy of 95%.

Senapaty et al. [23] introduced an IoT-enabled soil nutrient classification and crop recommendation (IoTSNA-CR) model to assist farmers throughout the cultivation process. The model employs sensors to gather real-time data on soil moisture, temperature, water, and nutrient levels. This work is comparable to Bhola et al. [25], where the model helps in optimal crop selection and reduces fertilizer usage. The innovative hybrid algorithm, MSVM-DAG-FFO, which combines a multi-class SVM with directed acyclic graph optimization and fruit fly optimization, achieved a remarkable 97.3% accuracy, surpassing the SVM and Decision Tree performance. Swaminathan et al. [24] address the challenge of fertilizer recommendations based on soil nutrients through a nutrient-centered deep collaborative filtering approach. This research aims to create an advanced recommender system called the Nutrient-centered Deep Collaborative Filtering (NDCF) method. The proposed method achieved root mean square and mean absolute error of 0.8411 and 0.655 on the collected dataset, respectively. The study by Khan et al. [26] presents a real-time context-aware fertilizer recommendation system utilizing ML and IoT technology. The model suggests suitable fertilizers for specific soil and crop types by capturing real-time soil fertility context through IoT-assisted mapping. The proposed IoT-assisted fertility mapping aligns well with standard soil chemical analysis, demonstrating mean differences of 0.34, 0.36, and -0.13 for Nitrogen, Phosphorous, and Potassium, respectively. Machine learning models are employed for context-aware fertilizer recommendation, including Logistic Regression, Support Vector Machine, Gaussian Naïve Bayes, and K-Nearest Neighbor. The Gaussian Naïve Bayes model exhibited the highest accuracy, reaching 96% and 94% for training and testing datasets.

Further, integrating modern technologies like the IoT and Artificial Intelligence (AI) in agriculture is crucial for efficient and quality crop production. Swaminathan et al. [27] introduce a four-layer architectural model encompassing sensors, networks, services, and applications to establish an energy-efficient farming system. The application layer employs deep learning for a fertilizer recommendation system aligned with expert opinions. The entire system is accessible through a user-friendly mobile application for farmers. This work underscores the significance of IoT-driven agricultural sensors and AI applications to enhance crop yield and sustainability, addressing the increasing demands of a growing population. The proposed approach showcases its potential to improve crop yield through fertilizer recommendations based on chemical properties. It includes integrating more sensors for comprehensive farm management. Thorat et al. [34] focus on integrating AI and sensor technology in agriculture to enhance insecticide, fertilizer recommendation, and soil nutrient analysis. The proposed approach employs Transition Probability Function (TPF) and Convolutional Neural Network (CNN), achieving over 90% accuracy. Indeed, the suitability of a particular soil type for various crops is significant, but unfavorable crop selection in a given location can negatively impact crop yield. This study addresses the identified limitations by introducing suitable crop selection and crop-specific fertilizer prediction architecture.

3. Methodology. Farmers are grappling with reduced crop yields and profits due to a limited understanding of crop selection intricacies and the factors influencing crop growth. Given that crop selection and the corresponding fertilizer stand as pivotal factors for maximizing yield and profitability, the primary objective of this study is to devise a Crop selection and Fertilizer recommendation framework to enhance agricultural returns.

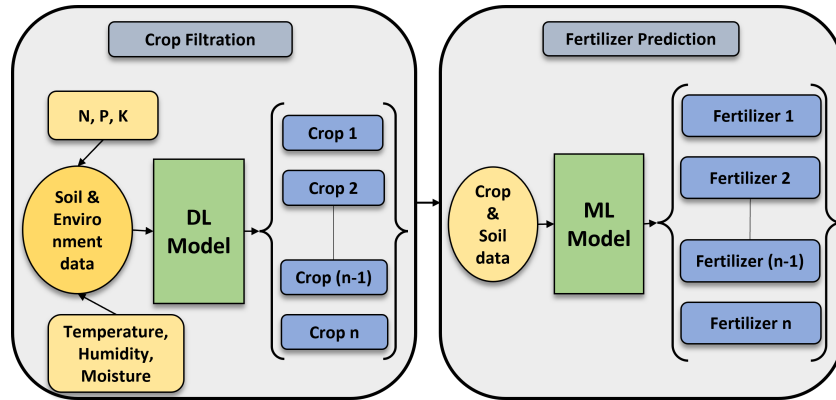


Fig. 3.1: Proposed framework

Consider a set of ‘ i ’ distinct crops denoted as $\{C_1, C_2, \dots, C_i\}$ and a collection of ‘ j ’ diverse farmlands denoted as $\{F_1, F_2, \dots, F_j\}$. It is assumed that each farmer possesses a Soil Health Card (SHC) for their farmland ‘ F_j ’, containing information about soil nutrient levels, alongside regular meteorological updates provided by government agencies. The objective is to determine appropriate crops and their corresponding fertilizers for each farmland based on inputs related to soil and weather conditions. The proposed framework, as depicted in Figure 3.1, employs a two-phase approach to suggest a variety of crops and associated fertilizers for each farm. In the initial phase, ‘ n ’ crops are filtered for each farmland ‘ F_j ’ from the pool of ‘ i ’ different crops. This phase involves assessing the compatibility of available crops with the local soil and weather conditions and filtering suitable crops. Further, the filtered crops are directed to the next stage, which identifies the suitable fertilizer for each crop for the farmer’s land. Implementing fertilizer tailored to the specific site offers benefits in terms of environmental sustainability, economic efficiency, and enhanced yield [35]. Each of these phases is further elaborated in the following subsections.

3.1. Crop filtration. Figure 3.2 illustrates the initial phase that focuses on filtering ‘ n ’ suitable crops. For every farmland ‘ F_j ’, consider $\{W_1(t), W_2(t), \dots, W_k(t)\}$ as weather conditions like temperature, rainfall, etc., at a time ‘ t ’, and $\{S_1(t), S_2(t), \dots, S_l(t)\}$ as soil attributes such as N, P, K, etc. Meteorological departments or local government agencies provide regular weather updates ‘ $W_k(t)$ ’ to farmers, aiding them in making informed decisions for their farming activities. Furthermore, the government provides a soil health card containing 12 vital soil macro and micro nutrients ‘ $S_l(t)$ ’, including pH, EC, Organic Carbon (OC), Nitrogen (N), Phosphorus (P), Potassium (K), Sulphur (S), Zinc (Zn), Boron (B), Iron (Fe), Manganese (Mn), and Copper (Cu). Since crop growth is intrinsically influenced by weather and soil conditions, these critical soil parameters are retrieved from the farmer’s soil health card and government weather updates to determine the most suitable crops. A proposed deep learning model computes probabilities $\{p_1, p_2, \dots, p_x\}$ based on these input parameters to rank crops. Subsequently, the top ‘ n ’ crops are selected and passed on to the next phase for further estimation.

Figure 3.4 portrays the architecture of the proposed artificial neural network model employed in the first phase. In this feed-forward backpropagation network, elements like weights, biases, the number of hidden layers, hidden neurons, learning rate, and training epochs play a pivotal role in determining prediction accuracy. These parameters are selected for precise predictions through a trial-and-error approach. A total of 7 inputs are fed into the input layer, and along with the bias, they are propagated to the hidden layer. The activation function ReLU is implemented for the hidden layers, while the output layers utilize the softmax activation function to predict probabilities. Furthermore, each hidden layer encompasses 512 neurons, the input layer consists of 7 neurons, and the output layer comprises 22 neurons, corresponding to each crop.

3.2. Fertilizer prediction. The second phase of the framework involves predicting fertilizer for each of the ‘ n ’ filtered crops obtained from the first phase. Figure 3.3 depicts the second phase of the ML-CSFR framework that predicts fertilizer using an ML model for each filtered crop individually on the available farmer’s

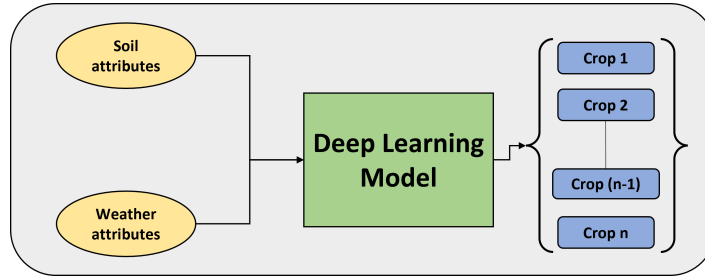


Fig. 3.2: Crop Filtration phase

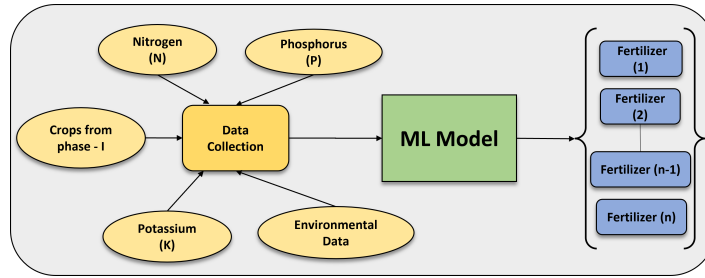


Fig. 3.3: Filter prediction phase

land. For every farmland ‘ F_j ’, filtered crops $\{C_1, C_2, \dots, C_i\}$, weather conditions $\{W_1(t), W_2(t), \dots, W_k(t)\}$, soil attributes $\{S_1(t), S_2(t), \dots, S_l(t)\}$, and $\{Fr_1, Fr_2, \dots, Fr_m\}$ be the available fertilizers.

The goal of the model is to recommend the appropriate fertilizer (y) for each filtered crop from the input set ‘ X ’, as described by Equation 3.1. The soil fertility level is determined by the quantities of nitrogen, phosphorus, and potassium in the soil. Equation 3.2 and Equation 3.3 represent the input and output feature set.

$$y = f(X) \quad (3.1)$$

$$X = \{W_k(t), S_l(t), C_m(t)\} \quad (3.2)$$

$$Y = F_n(t) \quad (3.3)$$

where, ‘ X ’ is the variables set representing weather, soil, and crop parameters; ‘ Y ’ represents the collection of commercially available fertilizers utilized as the model’s output for every combination of input variables. Each fertilizer available in the market possesses distinct compositions of essential nutrients like Nitrogen (N), Phosphorus (P), and Potassium (K), typically indicated in their names. The model inputs filtered crops, weather, and soil parameters and predicts suitable fertilizers for the land. Various models, including LR, DT, GaNB, LrSVM, XGB, and RF, are employed and compared to determine the most effective classification model for this phase.

3.3. Machine Learning models. Various machine learning models were employed and compared to determine the most effective classification model for fertilizer prediction. The models used include Logistic Regression, Decision Tree, Gaussian Naive Bayes, Support Vector Machine, Random Forest, Gradient Boosting, and XGBoost. Below is a detailed description of the setup for each model:

3.3.1. Logistic Regression. Logistic Regression is a linear model for binary classification. It models the probability that a given input point belongs to a particular class. The probability is modeled using a logistic

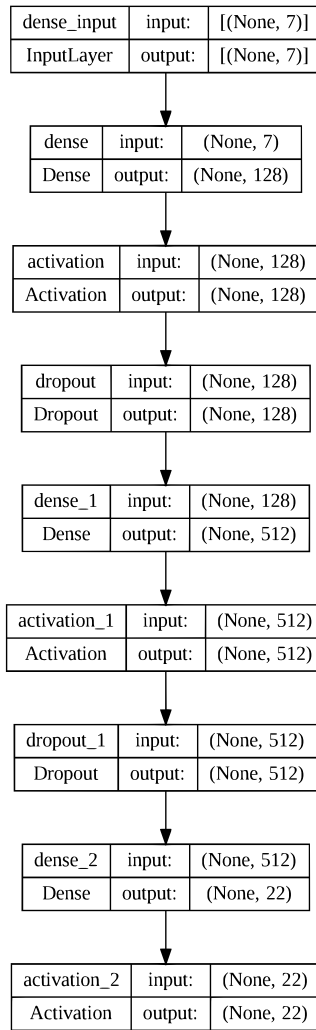


Fig. 3.4: Proposed ANN Architecture for Crop Filtration phase

function:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}} \tag{3.4}$$

where $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$ are the parameters to be estimated.

3.3.2. Decision Tree. Decision Trees are non-parametric models that split the data based on feature values. Each internal node represents a "test" on an attribute, each branch represents the outcome of the test, and each leaf node represents a class label. The Gini impurity or information gain is often used as a criterion to split the nodes:

$$Gini(D) = 1 - \sum_{i=1}^C P_i^2 \tag{3.5}$$

where P_i is the probability of an element being classified to a particular class.

3.3.3. Gaussian Naive Bayes. Gaussian Naive Bayes is based on Bayes' theorem and assumes independence between features. For continuous features, it assumes a Gaussian distribution:

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}} \quad (3.6)$$

where μ_y and σ_y^2 are the mean and variance of the feature x_i for class y .

3.3.4. Support Vector Machine. SVMs find the hyperplane that best separates the classes by maximizing the margin between the nearest points of the classes (support vectors). For the linear case, the decision function is:

$$f(x) = \beta_0 + \sum_{i=1}^n \beta_i x_i \quad (3.7)$$

with the optimization objective:

$$\min_{\beta} \frac{1}{2} \|\beta\|^2 \quad \text{subject to} \quad y_i(\beta_0 + \sum_{j=1}^n \beta_j x_{ij}) \geq 1 \quad \forall i \quad (3.8)$$

3.3.5. Random Forest. Random Forest is an ensemble method that constructs multiple decision trees during training and outputs the mode of the classes (classification) of the individual trees. The model's prediction is the average prediction of the individual trees. The construction process involves:

$$\hat{f}(x) = \frac{1}{B} \sum_{b=1}^B f_b(x) \quad (3.9)$$

where B is the number of trees and $f_b(x)$ is the prediction of the b_{th} tree.

3.3.6. Gradient Boosting. Gradient Boosting builds trees sequentially, with each new tree aiming to correct the errors of the previous one. The key idea is to fit a new model to the residual errors made by the previous model:

$$F_m(x) = F_{m-1}(x) + \lambda h_m(x) \quad (3.10)$$

where $F_m(x)$ is the ensemble model at stage m , $h_m(x)$ is the new tree, and λ is the learning rate.

3.3.7. XGBoost. XGBoost is an advanced implementation of gradient boosting with additional regularization terms to control overfitting. The objective function is:

$$L(\phi) = \sum_{i=1}^n l(\hat{y}_i, y_i) + \sum_{k=1}^K \Omega(f_k) \quad (3.11)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \quad (3.12)$$

where $\Omega(f_k)$ is a regularization term, T is the number of leaves, w is the leaf weight, and γ and λ are regularization parameters.

4. Experiment. This section provides a comprehensive evaluation of the proposed architecture through empirical analysis. It begins by outlining the experimental setup, detailing the dataset source, and discussing data analysis methods. Subsequently, the execution of the suggested model is discussed, followed by a comparative assessment of the attained outcomes compared to alternative machine learning models.

Table 4.1: Feature description (Crop filtration phase)

Features	Description	Unit
Nitrogen (N)	It is responsible for photosynthesis in the plant.	kg/ha
Phosphorus (P)	It is crucial to the crop's development.	kg/ha
Potassium (K)	It is required for the reproduction of crops.	kg/ha
pH level (pH)	It determines the availability of essential plant nutrients.	pH value
Temperature	Temperature is a key factor in plant growth and development.	degree Celsius
Humidity	Humidity is important for photosynthesis in plants.	%
Rainfall	The primary source of water for agricultural production.	mm

Table 4.2: Feature description (Fertilizer prediction phase)

Feature(s)	Description
Temperature	Degree Celsius
Humidity	%
Moisture	%
Soil type	5 types
Crop type	11 types
Nitrogen	Ratio
Potassium	Ratio
Phosphorous	Ratio
Fertilizer	7 types

4.1. Experimental setup. The experimental setup involves an Intel Core i5 processor with a quad-core x64-based architecture running at 3.6 GHz and 8 GB of RAM. The programming language used is Python, and the program was executed using the Google Colab notebook. Various standard software libraries such as Keras, Tensorflow, Matplotlib, and Numpy were utilized for implementation.

4.2. Dataset. Two distinct datasets are employed to assess the efficacy of the proposed ML-CSFR framework. The initial crop filtration phase dataset is sourced from [28]. This dataset categorizes lands and crops according to various attributes, encompassing soil characteristics such as nitrogen, phosphorus, potassium, and pH, and environmental factors impacting crop growth, such as humidity and rainfall. The collected dataset comprises 2200 land samples and encompasses 22 distinct crop types. Each crop category includes 100 individual land samples for analysis. The data has been divided into the ratio of 80:10:10 for training, validation, and testing sets. Table 4.1 describes the dataset features used in the first phase of the framework.

The second phase dataset [29] contains seven fertilizer varieties and eleven crops. The attributes of the collected dataset for the fertilizer dataset are shown in Table 4.2. The fertilizer used in executing the proposed approach is represented in Equation 4.1.

$$Y = \{Fr_1, Fr_2, \dots, Fr_m\} \quad (4.1)$$

The fertilizer names are used as ' Fr_1 ' for '20-20', ' Fr_2 ' for '14-35-14', ' Fr_3 ' for '26-28' fertilizer, ' Fr_4 ' for 'DAP', ' Fr_5 ' for '10-26-26', ' Fr_6 ' for 'Urea', and ' Fr_7 ' for '17-17-17'.

4.3. Dataset analysis. This section delves into the analysis of the soil and environmental data that influence both the crop filtration and fertilizer prediction processes. Initially, each dataset undergoes data cleaning procedures, as various attributes encompass distinct measurement scales. The Min-Max Scaler uses

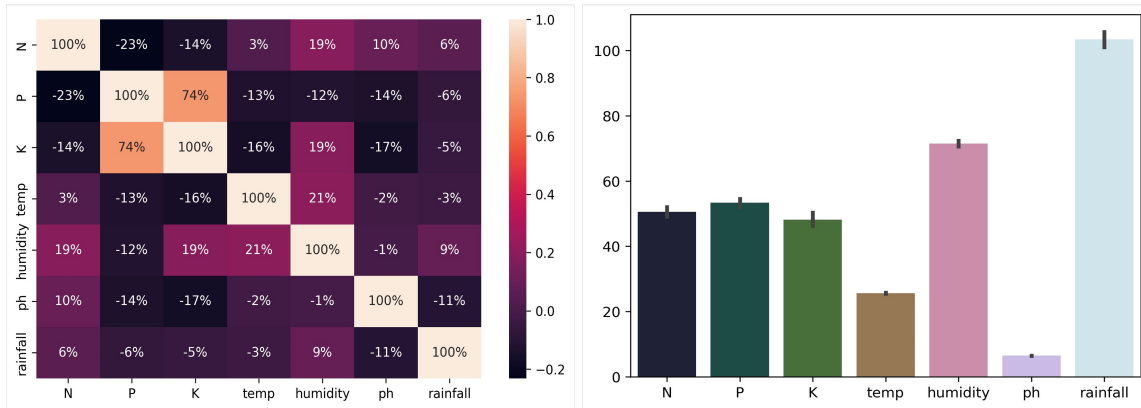


Fig. 4.1: Phase-1 dataset (a) Correlation matrix (b) Feature importance

the rescaled formula mentioned in Equation 4.2 for precise crop selection.

$$C = \frac{B - B_{min}}{B_{max} - B_{min}} \quad (4.2)$$

where C signifies the rescaled value, B is the feature value, B_{min} is the minimum value, and B_{max} is the maximum value. Further, the data has been apportioned into training and test sets, with an 80:10:10 split. The accuracy of these models is evaluated on both the training, validation, and test datasets.

The primary macronutrients, nitrogen, phosphorus, and potassium (N, P, and K), are pivotal in enhancing crop yield and quality. Figure 4.1(a) illustrates the correlation between the employed features, emphasizing the high correlation between potassium and phosphorous soil parameters and a moderate correlation between humidity and rainfall. On the other hand, Figure 4.1(b) provides insight into the crucial features within the crop filtration dataset, highlighting that among all the weather parameters, rainfall and humidity emerge as essential factors.

Figure 4.2 compares the nitrogen, phosphorus, and potassium values that different crops need. Among the crops, cotton, apple, and grapes exhibit the highest demand for macronutrients for optimal growth, whereas lentils, black gram, and oranges have the lowest requirements. The significance of various soil macronutrients such as N, P, and K holds a relatively uniform weightage across all crops. Notably, rainfall has the most significant importance among the considered parameters, while pH records are the least effective.

The dataset used for the study's second phase is sourced from [29]. The various commercial fertilizers used in the dataset and the distribution are classified in Figure 4.3. Many fertilizer products are labeled with the abbreviation NPK, followed by numerical values, such as NPK 10-26-26. The numbers following NPK indicate the percentage amounts of each nutrient in the fertilizer. For instance, an NPK value of 10-26-26 indicates that the fertilizer contains 10% nitrogen, 26% phosphorus, and 26% potassium [30]. The selected fertilizer and crop data type distribution in the dataset is visualized in Figure 4.4(a) and Figure 4.4(b), respectively.

Each type of soil has its unique attributes and composition. For instance, Urea is the most widely used solid nitrogen fertilizer and is usually applied as granules. Hence, its count is maximum in the dataset, as seen in Figure 4.4(a). Further, the crop has specific nutrient requirements for successful growth and enhanced production. Some crops necessitate lower nutrient levels, while others demand more. Consequently, in addition to soil type, crop type significantly influences fertilizer needs and recommendations.

The dataset distribution of the three macronutrients is categorized based on the provided classification. The application of fertilizer heavily relies on factors such as crop type, soil, and soil fertility regarding NPK nutrients. The soil nutrient level is interconnected with crops and fertilizers needed. Moreover, the relationships between crop type, soil type, and existing nutrient levels are even more intricate.

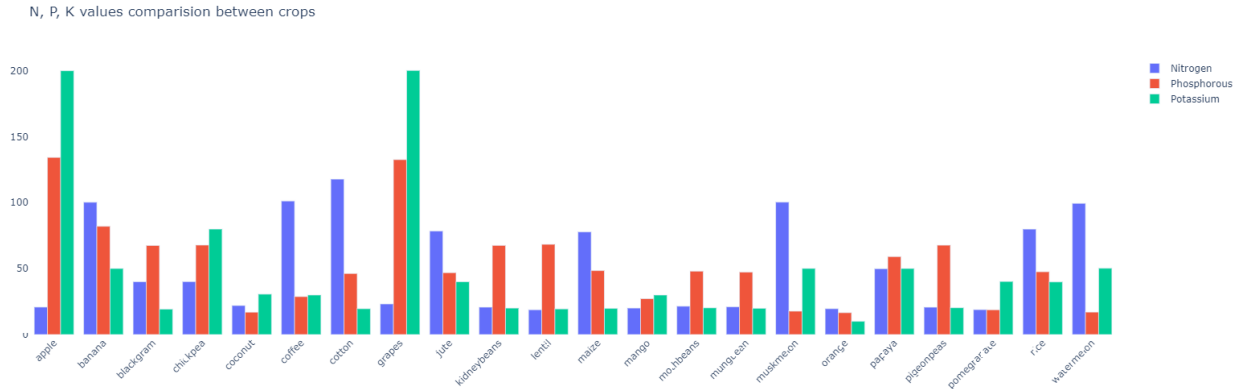


Fig. 4.2: N, P, K values required by different crops

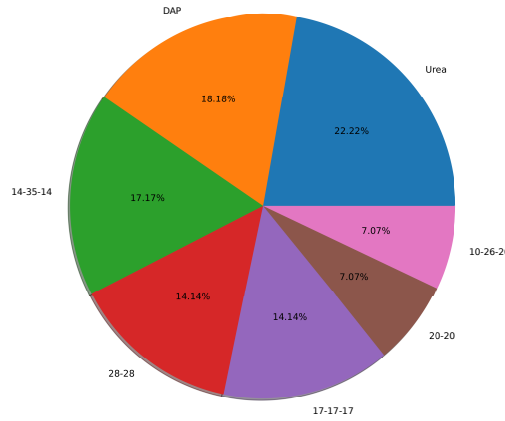


Fig. 4.3: Fertilizer class distribution

4.4. Algorithm for ML-CSFR framework. The primary goal of this experiment is to develop a recommendation model that will advise crop filtration and crop-specific fertilizer on various factors such as soil constituents, crop traits, and climate. Algorithm 1 presents the algorithm with the detailed steps involved in crop-based fertilizer prediction using ANN-XGB.

The algorithm is divided into two parts: (1) compute each crop’s rank and filter the top-n crops; (2) predict fertilizer for each crop corresponding to the farmer’s land. It requires soil health card details and environmental values concerning each land as input.

4.5. Evaluation Metrics. This study involves two separate datasets corresponding to the two phases of the framework. In the initial phase, the framework filters crops from multiple crop categories, while in the second phase, it predicts fertilizers from a range of different classes. As a result, a multi-class classification approach is utilized, and a confusion matrix is generated to calculate classification instances, including True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). In the context of multi-class classification, these instances can be interpreted as follows:

- True Positive (TP): The model correctly predicts the positive class.
- False Positive (FP): The model incorrectly predicts the positive class.
- True Negative (TN): The model correctly predicts the negative class.
- False Negative (FN): The model incorrectly predicts the negative class.

Algorithm 1 ML-CSFR: Crop Selection and Fertilizer Recommendation**Require:** Local input variables $X = \{x_1, x_2, \dots, x_n\}$, Soil parameters S , Weather parameters W **Ensure:** Top 3 recommended crops $C = \{c_1, c_2, c_3\}$, Recommended fertilizers F

```

1: Phase 1: Crop Selection using ANN
2: Initialize ANN parameters: number of layers  $L$ , neurons in each layer, activation functions
3: for epoch = 1 to  $N$  do
4:   for each batch in data do
5:     Forward Propagation:
6:     for layer  $l = 1$  to  $L$  do
7:        $z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)}$ 
8:        $a^{(l)} = \sigma(z^{(l)})$ 
9:     end for
10:    Output Layer:
11:     $\hat{y}_i = \frac{e^{a_i^{(L)}}}{\sum_{j=1}^m e^{a_j^{(L)}}}$ 
12:    Loss Function:
13:     $\mathcal{L} = -\sum_{i=1}^m y_i \log(\hat{y}_i)$ 
14:    Backward Propagation:
15:    Compute gradients:  $\frac{\partial \mathcal{L}}{\partial W^{(l)}}$  and  $\frac{\partial \mathcal{L}}{\partial b^{(l)}}$ 
16:    Update weights and biases:
17:     $W^{(l)} := W^{(l)} - \eta \frac{\partial \mathcal{L}}{\partial W^{(l)}}$ 
18:     $b^{(l)} := b^{(l)} - \eta \frac{\partial \mathcal{L}}{\partial b^{(l)}}$ 
19:    end for
20:  end for
21: Output top 3 recommended crops  $C = \{c_1, c_2, c_3\}$  based on highest probabilities  $\hat{y}_i$ 
22: Phase 2: Fertilizer Recommendation using XGBoost
23: Initialize XGBoost parameters: number of trees  $T$ , learning rate  $\eta$ , maximum depth  $d$ 
24: for  $t = 1$  to  $T$  do
25:   Compute predictions  $\hat{y}_t = \sum_{i=1}^t f_i(X)$ 
26:   Compute residuals  $r_t = y - \hat{y}_t$ 
27:   Fit regression tree to residuals:  $f_t = \arg \min_f \sum_{i=1}^n L(y_i, \hat{y}_{t-1} + f(x_i))$ 
28:   Update model:  $\hat{y}_t := \hat{y}_{t-1} + \eta f_t(X)$ 
29: end for
30: For each crop  $c \in C$ , predict the suitable fertilizer  $F$  using XGBoost
31: Output: Recommended crops  $C$  and fertilizers  $F$ 

```

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (4.3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4.4)$$

$$Recall = \frac{TP}{TP + FN} \quad (4.5)$$

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4.6)$$

Equation 4.3 represents accuracy, calculating the ratio of correctly predicted observations to total observations. Equation 4.4 defines precision, which measures the ratio of the true positive predictions to the total number of positive predictions. In contrast, Equation 4.5 represents recall, which determines the number of true positive predictions divided by the total number of relevant observations. The F_1 Score, as calculated in Equation 4.6, represents the harmonic mean of precision and recall.

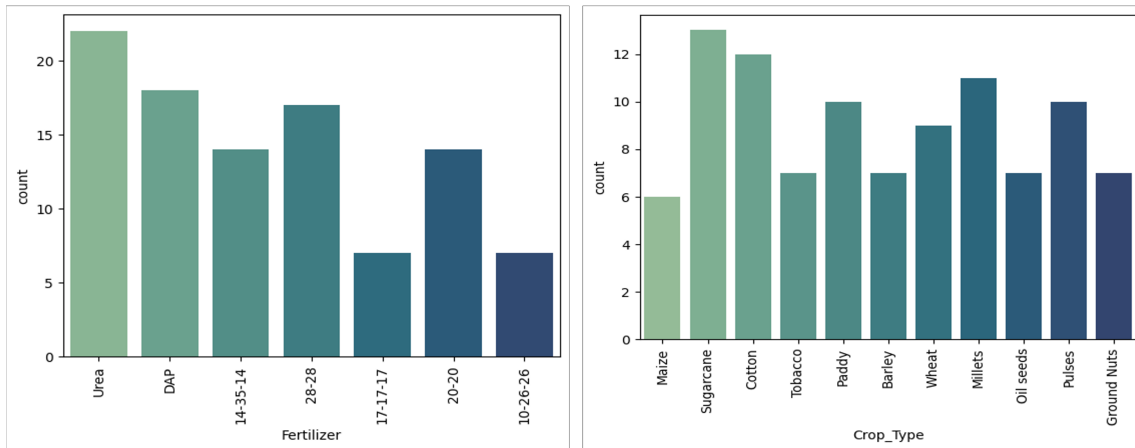


Fig. 4.4: Fertilizer data distribution (a) Fertilizer type (b) Crop type

Table 5.1: Comparative analysis (Crop filtration phase)

Accuracy / Models	DT	SVM	RF	KNN	XGBoost	ANN
Training Accuracy	0.985	0.985	0.992	0.988	0.992	0.992
Testing Accuracy	0.974	0.976	0.985	0.976	0.981	0.991
Validation Accuracy	0.972	0.975	0.983	0.975	0.980	0.990
Precision	0.972	0.975	0.983	0.975	0.981	0.991
Recall	0.974	0.976	0.985	0.976	0.982	0.992

Table 5.2: List of filtered crops for sample lands

Land	Top filtered crops
Land 1	Rice, Maize, Cotton
Land 2	Maize, Rice, Jute
Land 3	Maize, Chickpea, Lentil

5. Results. A series of comparative experiments were conducted to evaluate the performance of the proposed framework for crop filtration and crop-specific fertilizer recommendation. The results of these experiments are comprehensively analyzed and discussed in this section.

5.1. Result Analysis of Crop Filtration Phase. The crop filtration phase utilizes a classification model to filter crops based on their probabilities. The obtained results for the initial phase under different environmental conditions are summarized in Table 5.1.

The evaluation outcomes demonstrate that the proposed Artificial Neural Network (ANN) model achieves a validation accuracy of 99.07%, testing accuracy of 99.10%, precision of 99.13%, and recall of 99.24%. These results are the highest among all the studied models. Additionally, the Random Forest (RF) and XGBoost (XGB) models also performed well, with testing accuracies, precision, and recall all above 98%. Conversely, the Decision Tree (DT) recorded the lowest validation and testing accuracies of 97.20% and 98.50%, respectively. This lower accuracy could be attributed to variations in the dataset, as DT is very sensitive to small perturbations in the data.

Table 5.2 provides an overview of the crops the ANN model filtered for three distinct farmlands. While the present study filters only three crops, the number of crops can be customized according to the user's preferences.

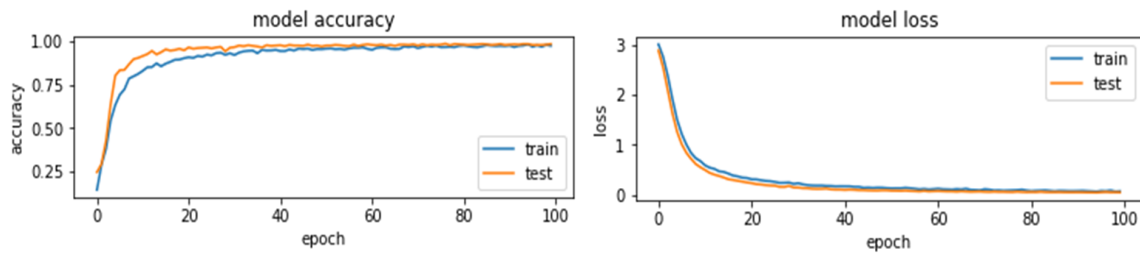


Fig. 5.1: ANN result (a) accuracy vs epoch (b) loss vs epoch

Table 5.3: Performance Comparison (Fertilizer Prediction Phase)

Model	Training Accuracy	Testing Accuracy	Precision	Recall	F1-Score	AUC
LR	0.92	0.70	0.71	0.70	0.70	0.75
DT	0.99	0.94	0.94	0.94	0.94	0.96
GaNB	0.96	0.50	0.52	0.50	0.49	0.55
LrSVM	0.94	0.85	0.86	0.85	0.85	0.88
XGB	0.99	0.97	0.97	0.97	0.97	0.98
RF	0.99	0.95	0.95	0.95	0.95	0.97

Figures 5.1(a) and (b) illustrate the graphs depicting accuracy versus epoch and loss versus epoch for the proposed Artificial Neural Network (ANN) model. Examining the loss curve reveals that the global optimal minimum is attained in the initial stages of iterations. The results justify the selection of ANN as the proposed model for crop filtration as its performance consistently improves and error reduces over time.

5.2. Result Analysis of Fertilizer Prediction Phase. In the fertilizer prediction phase, the performance of various machine learning models is evaluated, and the most effective model is selected for this phase. The most appropriate model is chosen by considering the performance metrics such as accuracy, precision, recall, and F1-score. The models, including LR, DT, GaNB, LrSVM, XGB, and RF, were applied to the preprocessed dataset for fertilizer prediction. The performance assessment of these models is detailed in Table 5.3.

The analysis revealed that tree-based models like RF, XGB, and DT exhibited more exceptional performance stability than other models. This stability arises from the fact that these models establish decision boundaries that are more consistent and accurate by aggregating the outcomes of multiple trees and utilizing majority voting to arrive at a precise prediction.

Notably, the XGB performs exceptionally well in recommending suitable fertilizers, achieving an accuracy of 99.23% and 97.66% on the training and test datasets, respectively. The classification report and confusion matrix for the best-performing model XGB are shown in Table 5.4 and Figure 5.2, respectively. The table indicates that all the fertilizers achieved 1.00 precision except 17-17-17, which achieved a precision of 0.67. Further, fertilizer 10-26-26 achieves a low recall of 0.67, indicating that the proposed model misclassifies 33 out of 100 classes. The low recall value can be attributed to the imbalance dataset, as depicted in Figure 4.3, where class 10-26-26 only accounts for 7.07% of the data.

Conversely, the GaNB model is the worst-performing, primarily due to its unsuitability for dealing with imbalanced class problems. When dealing with imbalanced datasets, where certain classes have much larger instances than others, GaNB can encounter difficulties delivering accurate outcomes. This problem may lead to biased predictions favoring the majority class and suboptimal performance for the minority class.

In summary, XGB performed better than other models in testing accuracy, precision, recall, F1-score, and AUC values. Table 5.4 present the classification report of the XGB model. The table depict a consistent performance in model's accuracy, with a reduced error over time, thereby justifying the selection of XGB as

Table 5.4: Classification report (XGB)

Classes	Code	Precision (%)	Recall (%)	F1-Score
10-26-26	0	1.00	0.67	0.80
14-35-14	1	1.00	1.00	1.00
17-17-17	2	0.67	1.00	0.80
20-20	3	1.00	1.00	1.00
28-28	4	1.00	1.00	1.00
DAP	5	1.00	1.00	1.00
Urea	6	1.00	1.00	1.00
Accuracy				0.97

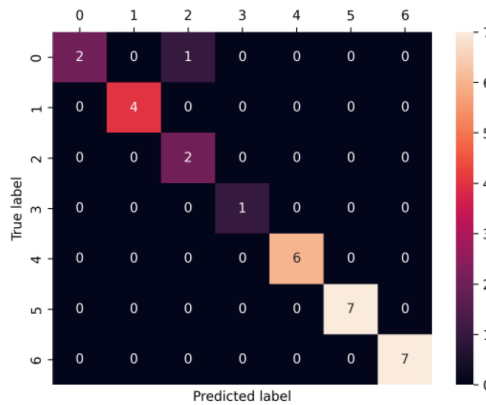


Fig. 5.2: Confusion matrix (XGB)

the proposed fertilizer recommendation model.

The fertilizer recommendations for each of the filtered crops for Land 1 are outlined in Table 5.5. The proposed model suggested DAP, 20-20, and 14-35-14 as the appropriate fertilizers for the predicted crops: Rice, Maize, and Cotton, respectively. These recommendations are substantiated by the references provided in the table. For example, the prediction of DAP or Urea fertilizer for the Rice crop aligns with established practices [31]. Moreover, the 28-28 fertilizer, a complex blend rich in Nitrogen and Phosphorus, is suitable for crops like Paddy, Cotton, Chillies, Sugarcane, and Vegetables, offering immediate and sustained greenness [32]. Similarly, the 14-35-14 fertilizer is optimal for Rice, Cotton, Groundnut, Chillies, Soya bean, and Potato crops, particularly those demanding high initial Phosphate [33]. Thus, the proposed framework is validated by its ability to suggest fitting crop-specific fertilizers, providing farmers with insights to enhance yields.

5.3. Discussion. The development of the proposed two-phase framework for crop filtration and fertilizer recommendation represents a significant advancement in precision agriculture. The experimental results underscore the framework's effectiveness, showcasing its ability to make accurate and reliable recommendations tailored to specific agricultural conditions.

In the crop filtration phase, the Artificial Neural Network (ANN) model demonstrated superior performance compared to other classification models, achieving a remarkable testing accuracy of 99.10%, precision of 99.13%, and recall of 99.24%. These metrics indicate the model's robustness in correctly identifying suitable crops under varying environmental conditions. The high precision and recall values further affirm the model's ability to make precise and consistent predictions, minimizing false positives and negatives.

The fertilizer prediction phase involved evaluating various machine learning models, where XGBoost (XGB) emerged as the best-performing model. It achieved training and testing accuracies of 99.23% and 97.66%, respec-

tively, along with a high F1-score and AUC, indicating its exceptional ability to distinguish between different classes of fertilizers. The classification report for XGB showed that most fertilizers achieved precision and recall of 1.00, except for fertilizer 17-17-17 and 10-26-26, which had a lower precision and recall. This discrepancy can be primarily attributed to class imbalance, as evidenced by the confusion matrix and distribution of the dataset.

Further, the framework iteratively follows a dynamic update mechanism, continuously updating its recommendations based on the latest soil and weather data. After each planting season, soil health is reassessed using an updated Soil Health Card or periodic soil testing. This updated soil data and ongoing weather information are fed back into the framework, which reruns the crop and fertilizer recommendation process. This dynamic feedback loop ensures that the recommendations remain relevant and accurate over time, adapting to changes in soil nutrient levels resulting from fertilizer applications and other environmental factors.

5.3.1. Error-Analysis. The error analysis from the first phase revealed that the Decision Tree (DT) model recorded the lowest validation and testing accuracies of 97.20% and 98.5%, respectively. This lower performance can be attributed to the model's sensitivity to small perturbations in the dataset, highlighting its limitations in handling imbalance data. Conversely, the Random Forest (RF) and XGBoost (XGB) models also performed well, with testing accuracies above 98%, but were slightly outperformed by the ANN model, suggesting that the latter's deeper architecture and non-linear learning capabilities offer significant advantages in this application.

Further, from the second phase error analysis is the model's occasional misclassification of minority class fertilizers, such as 10-26-26, which had a recall of 0.67. This misclassification suggests further data balancing techniques or augmenting the minority class samples to improve the model's performance on underrepresented classes.

5.3.2. Contributions. The significant contributions of this work include:

- *High Accuracy and Reliability:* The ANN model achieved exceptional accuracy in the crop filtration phase, and the XGB model demonstrated superior performance in the fertilizer prediction phase. These results highlight the models' effectiveness in handling complex agricultural data and making accurate recommendations.
- *Robustness to Environmental Variations:* The high precision and recall values across different models indicate the framework's robustness and adaptability to varying environmental conditions, ensuring reliable performance in diverse agricultural settings.
- *Scalability and Practicality:* The framework's architecture is designed to be scalable and capable of processing extensive agricultural data inputs in a cloud-based environment. This scalability ensures its applicability to small-scale farms and large agricultural enterprises.
- *Real-world Applicability:* The proposed framework's ability to provide farm-specific recommendations and its potential integration with real-time sensor data make it a practical tool for enhancing decision-making in agriculture. The accurate predictions of crop selection and fertilizer recommendations can lead to improved crop yields and optimized resource use.

Further, the proposed ML-CSFR framework is highly relevant to scalable computing due to its inherent ability to handle real-life problems. Using machine learning models such as ANN and XGBoost allows the framework to process extensive agricultural data inputs, including soil characteristics and weather conditions, in a scalable manner. This scalability ensures that the framework can be applied to diverse agricultural settings, from small-scale farms to large agricultural enterprises, that can be a versatile solution for improving crop selection and fertilizer recommendation processes. Additionally, the framework's architecture is designed to be deployable on cloud-based platforms, leveraging the power of distributed computing further to enhance its scalability and applicability in different agricultural scenarios.

Finally, the significant observation drawn from the results is the remarkable effectiveness of the proposed model in scenarios demanding farm-specific recommendations. These findings further confirm the practicality of the proposed framework, particularly in situations characterized by limited resources and constraints. Moreover, besides its simple architecture, the framework serves a dual role of crop selection and associated fertilizer recommendation, eliminating the requirement for separate applications for these tasks. This reduction in

Table 5.5: Recommended Fertilizer

Crop	Recommended Fertilizer	Ref.
Rice	DAP	[31]
Maize	28-28	[32]
Cotton	14-35-14	[33]

overhead enhances the model's efficiency and practicality. These findings underscore the adaptability of the proposed framework, positioning it as a valuable tool ready for implementation in real-world scenarios.

6. Conclusions and Future Works. The exponential growth of the world's population has boosted demand for both quantity and quality of food. Consequently, the agricultural sector is required to undergo a modern transformation to address this demand adequately. The integration of modern technologies with intelligent algorithms holds the potential to benefit the farming community significantly. This paper proposes ML-CSFR, a two-phase Crop Selection and Fertilizer Recommendation framework designed to provide better returns for the agricultural sector. The framework is a machine-learning tool that provides crop selection and fertilizer recommendations by leveraging local input variables. It harnesses various weather and soil data sources to deliver precise recommendations. The initial phase of crop filtration is executed using Artificial Neural Networks. The result justifies that ANN outperforms other ML models with an accuracy of 99.10%. In the second phase of fertilizer prediction, the results indicate that XGBoost is the most accurate machine learning model, achieving an accuracy of 99.23% for the training dataset and 97.66% for the test dataset. The proposed work contributes to increased accuracy in crop prediction, improved soil health with proper fertilization, and enhanced decision-making.

Additionally, potential areas for future exploration involve enlarging the datasets used and extending the application's utility to gain more comprehensive insights into crop and soil management. In terms of practical implications, the simple and lightweight design of the suggested framework offers potential for future integration with handheld devices. This integration could help farmers conveniently forecast crops and their corresponding fertilizers. Moreover, enhancing recommendations could involve integrating real-time sensor data collected from crop fields. This approach would enable precise determination of the exact quantities of macro and micro nutrients required, tailored to the specific demands of each crop.

REFERENCES

- [1] Agriculture & Farmers Welfare, M. Contribution of Agriculture Sector in GDP. (<https://www.pib.gov.in/PressReleasePage.aspx?PRID=1909213>), Accessed: 2023-08-03
- [2] Aayog, N. Workforce Changes and Employment. (https://www.niti.gov.in/sites/default/files/2023-02/Discussion_Paper_on_Workforce_05042022.pdf), Accessed: 2023-08-03
- [3] Food & United Nations, A. Food wastage: Key facts and figures. (<https://www.fao.org/news/story/en/item/196402/icode>), Accessed: 2023-08-03
- [4] Veeragandham, S. & Santhi, H. A review on the role of machine learning in agriculture. *Scalable Computing: Practice And Experience*. **21**, 583-589 (2020)
- [5] Bhola, A. & Kumar, P. Performance Evaluation of Different Machine Learning Models in Crop Selection. *Robotics, Control And Computer Vision*. pp. 207-217 (2023)
- [6] Verma, S., Kumar, P. & Singh, J. A Unified Lightweight CNN-based Model for Disease Detection and Identification in Corn, Rice, and Wheat. *IETE Journal Of Research*. pp. 1-12 (2023)
- [7] Bhola, A., Verma, S. & Kumar, P. A comparative analysis of deep learning models for cucumber disease classification using transfer learning. *Journal Of Current Science And Technology*. **13**, 23-35 (2023)
- [8] Sinha, A., Shrivastava, G. & Kumar, P. Architecting user-centric internet of things for smart agriculture. *Sustainable Computing: Informatics And Systems*. **23** pp. 88-102 (2019)
- [9] Swaminathan, B., Palani, S., Vairavasundaram, S., Kotecha, K., & Kumar, V. IoT-driven artificial intelligence technique for fertilizer recommendation model. *IEEE Consumer Electronics Magazine*, **12** pp. 109-117. (2022)
- [10] Janvier, NIYITEGEKA and Arcade, N & Eric, NGABOYERA & Jean, N. Machine Learning based Soil Fertility Prediction. *International Journal of Innovative Science, Engineering & Technology*, **8** pp. 141-146. (2021)

- [11] Sujatha, M and Jaidhar, CD Machine learning-based approaches to enhance the soil fertility—A review. *Expert Systems with Applications*. pp. 122557 (2023)
- [12] Bhattacharya, S., & Pandey, M. PCFRIMDS: Smart Next-Generation Approach for Precision Crop and Fertilizer Recommendations Using Integrated Multimodal Data Fusion for Sustainable Agriculture. *IEEE Transactions on Consumer Electronics*. (2024)
- [13] Wu, M., Xiong, J., Li, R., Dong, A., Lv, C., Sun, D., ... & Niu, W. Precision forecasting of fertilizer components' concentrations in mixed variable-rate fertigation through machine learning. *Agricultural Water Management*. **298** pp. 108859. (2024)
- [14] Reddy, G. V., Reddy, M. V. K., Spandana, K., Subbarayudu, Y., Albawi, A., Chandrashekar, R., ... & Praveen, P. Precision farming practices with data-driven analysis and machine learning-based crop and fertiliser recommendation system. *In E3S Web of Conferences*. **507**, pp. 01078. (2024)
- [15] Kang, M., Wang, X., Wang, H., Hua, J., Reffye, P. & Wang, F. The development of AgriVerse: Past, present, and future. *IEEE Transactions On Systems, Man, And Cybernetics: Systems*. (2023)
- [16] Modi, A., Sharma, P., Saraswat, D. & Mehta, R. Review of Crop Yield Estimation using Machine Learning and Deep Learning Techniques. *Scalable Computing: Practice And Experience*. **23**, 59-80 (2022)
- [17] Jani, K. & Chaubey, N. A novel model for optimization of resource utilization in smart agriculture system using IoT (SMAIoT). *IEEE Internet Of Things Journal*. **9**, 11275-11282 (2021)
- [18] Cheema, S., Singh, A. & Gritli, H. Optimal Crop Selection Using Gravitational Search Algorithm. *Mathematical Problems In Engineering*. **2021** (2021)
- [19] Bakthavatchalam, K., Karthik, B., Thiruvengadam, V., Muthal, S., Jose, D., Kotecha, K. & Varadarajan, V. IoT framework for measurement and precision agriculture: predicting the crop using machine learning algorithms. *Technologies*. **10**, 13 (2022)
- [20] Jain, K. & Choudhary, N. Comparative analysis of machine learning techniques for predicting production capability of crop yield. *International Journal Of System Assurance Engineering And Management*. **13**, 583-593 (2022)
- [21] Gupta, R., Sharma, A., Garg, O., Modi, K., Kasim, S., Baharum, Z., Mahdin, H. & Mostafa, S. WB-CPI: Weather based crop prediction in India using big data analytics. *IEEE Access*. **9** pp. 137869-137885 (2021)
- [22] Mariammal, G., Suruliandi, A., Raja, S. & Poongothai, E. Prediction of land suitability for crop cultivation based on soil and environmental characteristics using modified recursive feature elimination technique with various classifiers. *IEEE Transactions On Computational Social Systems*. **8**, 1132-1142 (2021)
- [23] Senapaty, M., Ray, A. & Padhy, N. IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture. *Computers*. **12**, 61 (2023)
- [24] Swaminathan, B., Palani, S. & Vairavasundaram, S. Feature fusion based deep neural collaborative filtering model for fertilizer prediction. *Expert Systems With Applications*. **216** pp. 119441 (2023)
- [25] Bhola, A., & Kumar, P. Deep feature-support vector machine based hybrid model for multi-crop leaf disease identification in Corn, Rice, and Wheat. *Multimedia Tools and Applications*. pp. 1-21 (2024)
- [26] Khan, A., Faheem, M., Bashir, R., Wechtaisong, C. & Abbas, M. Internet of things (IoT) assisted context aware fertilizer recommendation. *IEEE Access*. **10** pp. 129505-129519 (2022)
- [27] Swaminathan, B., Palani, S., Vairavasundaram, S., Kotecha, K. & Kumar, V. IoT-driven artificial intelligence technique for fertilizer recommendation model. *IEEE Consumer Electronics Magazine*. **12**, 109-117 (2022)
- [28] INGLE, A. Crop Recommendation Dataset. (<https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset>), Accessed: 2023-07-03
- [29] Abhishek, G. Fertilizer Prediction Dataset. (<https://www.kaggle.com/datasets/gdabhishek/fertilizer-prediction>), Accessed: 2023-07-03
- [30] University, T. Nutrient Management: Fertilizers. (https://agritech.tnau.ac.in/agriculture/agri_nutrientmgt_fertilizers.html), Accessed: 2023-08-03
- [31] Agriculture, I. Fertilizer Application for Rice. (https://agri.punjab.gov.in/sites/default/files/fertilizer_application_for_rice.pdf), Accessed: 2023-08-03
- [32] IndiaMart GROMOR 28-28-0. (https://www.indiamart.com/proddetail/gromor-28-28-0-fertilizer-10329114088.html?smp_widget=1), Accessed: 2023-08-03
- [33] IndiaMart GROMOR 14-35-14. (<https://www.indiamart.com/proddetail/gromor-14-35-14-10329096730.html>), Accessed: 2023-08-03
- [34] Thorat, T., Patle, B. & Kashyap, S. Intelligent insecticide and fertilizer recommendation system based on TPF-CNN for smart farming. *Smart Agricultural Technology*. **3** pp. 100114 (2023)
- [35] Sanches, G., Magalhães, P., Kolln, O., Otto, R., Rodrigues Jr, F., Cardoso, T., Chagas, M. & Franco, H. Agronomic, economic, and environmental assessment of site-specific fertilizer management of Brazilian sugarcane fields. *Geoderma Regional*. **24** pp. e00360 (2021)
- [36] Vandana, W. M., & Kavya, B. Soil Fertility Assessment and Crop Recommendation for Sustainable Farming using Machine Learning and Deep Learning. *In 2024 4th International Conference on Data Engineering and Communication Systems (ICDECS)* pp. 1-3. IEEE. (2024)

Edited by: Katarzyna Wasielewska-Michniewska

Regular paper

Received: Sep 7, 2023

Accepted: Jul 29, 2024