SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I11P141

Volume 11 Issue 11, 448-460, November 2024

Original Article

A Novel Design and Implementation of Fertilizer Recommendation System Based on Hybrid Machine Learning Models

G. Mamatha[1]*, Jyothi S. Nayak[2]

[1]*Department of Computer Engineering and Technology, Chaitanya Bharathi Institute of Technology(A), Hyderabad, Telangana, India.*
[1,2]*Department of Computer Scienceand Engineering, B.M.S College of Engineering, Karnataka, India.*

**Corresponding Author : gmamatha_cet@cbit.ac.in*

Received: 24 September 2024 Revised: 25 October 2024 Accepted: 23 November 2024 Published: 30 November 2024

Abstract *- India is mostly a farming country. Many of the GDP of emerging nations like India comes from agriculture, so the sector is important to these countries' economies. The demand for food has skyrocketed due to the increase in population. Crop quality, yield, and profitability can all take a hit when farmers choose their crops, fertilizers, and pesticides without considering factors like soil type, water requirement, temperature conditions, and crop profitability analysis for a specific area. The development of computational technology has prompted scientists to consider various issues, including identifying diseases and pesticides and selecting fertilizers and crops based on soil quality, water needs, and market viability. An essential and fundamental aspect of farming is selecting the appropriate fertilizer for soil and crop production. India has a reputation as an agricultural powerhouse, with traditional practices still used to advise farmers on the best fertilizer to use. Communication between farmers and specialists is currently the basis for suggestions, and various experts have different recommendations. The prohibitive cost of lab technology for determining soil supplement levels is a major concern. Present frameworks for determining soil nutrient content and fertilizer recommendations are ineffective and inefficient. In order to estimate the nutritional dimension in soil and provide suitable fertilizer, this article presents an attractive, novel fertilizer recommendation system, 'FertRec'. In this proposed system, soil samples are analyzed to identify the deficiency of nutrients, thus preparing the datasets used for training the machine learning models. The best accuracy model recommends a suitable fertilizer based on the soil nutrients. The main goal is to create an effective fertilizer recommendation system to help farmers optimize their fertilizer use. Compared to current benchmark recommendation methods, the proposed system performs four times better on 500 soil samples from the Telangana region in India, using soil features, and it effectively recommends fertilizer with 99.98% accuracy. This will maximize production, and farmers greatly benefit from this method, which involves selecting the appropriate fertilizer at the beginning of the product cycle.*

Keywords *- Agriculture, Soil nutrients, Fertilizer recommendation, Machine learning models.*

1. Introduction

The agricultural sector is fundamental to the economy of any nation. Farmers have always relied on the information and advice of their peers to improve their farming practices, but this method is not always effective, and it can result in crop failure and financial losses. Conventional farming practices do not reveal precise information on soil qualities, precise water requirements for each crop, and other critical factors that prevent countries from meeting their food demands.

When it comes to a country's overall economic growth, the agriculture business is crucial. For farming to be successful, soil quality must be preserved. According to Casta neda-Miranda and Casta no-Meneses (2020) in [1], intelligent planning, analysis, and production control technologies are crucial for improving agriculture's organic soil productivity, plant nutrition, and water quality.

Soils used for agriculture facilitate plant growth by acting as a framework for exchanging water, nutrients, and gases. Plant growth is influenced by the soil's physical, chemical, and biological characteristics. "The capacity of soil to function" is the simplest way to describe soil quality. Soil quality is subjective and purpose-dependent, being a product of human engineering. Soil parameters vary naturally across landscapes, and when these variations impact environmental factors or agricultural yields, they become significant. Soil is an integral part of agricultural output. Therefore, taking care of it now will pay dividends in the long run. Soil water and energy management determines crop productivity without harming plants, soil, or the environment. It is necessary to provide farmers with instructions for the types and amounts of fertilizers to use depending on the findings of soil tests to keep the soil healthy and make farming more sustainable. Fertilizer recommendations for a calendar year's worth of crops should also be provided, considering various soil factors.

Agriculture plays a significant role in India's economy and has been seen as the driving force behind human advancement from ancient times. More than 60% of India's population relies on agriculture as their primary source of income, and the agricultural sector is an indisputable necessity for the country's Gross Domestic Product (GDP) growth. But these days, farmers are overwhelmed by the variety of fertilizers on the market and often employ the one that is most well-known in their area without considering it. Two big problems arise from this: low yield and soil contamination. Crop production is diminished because inadequate nutrients are applied during fertilizer application. Over fertilization contaminates both the land and the food grown there, leading to food sickness. Over fertilization causes soil acidity, root burn, mineral degradation, and groundwater contamination. For the harvest to flourish, it is crucial to identify these nutrients as early as possible. Soil nutrient estimation and characterization for fertilizer recommendation purposes is at the heart of the proposed system 'FertRec'.

Fertilization is crucial in agricultural production because it increases crop yield and quality by supplementing soil nutrients [2]. Commercial fertilizers can boost crop yields by 30–50% [3]. According to the literature, some small countries used more chemical fertilizers than any other large country in the past, and that number is only going up [4]. The problem of over-fertilization and improper methods of fertilization are ongoing issues in developing countries [5]. The amount of fertilizer used to plant crops does not directly correlate to the financial returns on investment. In addition to driving up the cost of agricultural production, excessive fertilizer use causes nutrient supply and crop nutrient demand to be inconsistent, making it difficult to enhance crop yields [6] and can even lead to decreases in crop output [7]. Agricultural non-point source pollution [8], soil fertility [9], and the sustainable increase of land productivity are all exacerbated by poor fertilization, which also leads to a considerable volume of fertilizer lost to the external environment.

Farmers currently send soil samples to nearby agricultural research centers for analysis. Additionally, the test results must be obtained from the identical facility after one week. Nevertheless, the human procedure of generating reports and providing them to farmers is laborious and time-consuming. Nevertheless, the local farmers are oblivious to mail access and report viewing basics**.** Fertilizers are compounds that provide plants with the nutrients they need to grow. Fertilizer is a crucial component of farming and is responsible for around 55% of the increase in production. Among the nutrients found in soil are ni-trogen (N), phosphorus (P), and potassium (K), with boron, iron, chlorine, copper, manganese, zinc, and nickel making up the minor nutrients [10]. Farmers apply fertilizers without knowing the soil's fertility. As nutrients build up in the soil, excess fertilizers can harm plants and lead to soil contamination. Because of this, sulphur and nitrogen oxides are released into the air. Eating heavy nitrogen-based green vegetables can negatively affect humans [11]. Improving agricultural output efficiency necessitates the establishment of a sophisticated fertilization decision system.

This article aims to create a novel fertilizer recommendation system that helps farmers make the best decisions about fertilizer usage to maximize their crop yields while cutting costs and increasing their profitability.

This study presents 'FertRec', a new recommendation system that relies on machine learning. The proposed system aims to suggest suitable fertilizer for each, according to predefined criteria. Soil type, soil qualities, water conditions, crop water requirements, land area, and crop market value were all factors considered by utilizing recommendation algorithms based on machine learning. This system analyzes soil test results to suggest fertilizers. The suggested approach takes crops and nutrients as inputs and outputs recommendations for fertilizers.

2. Related Literature Work

The fertilizer recommendation systems have been the subject of numerous studies. However, the applications are more intricate, and the farmers find them difficult to understand.

An intriguing side effect of using the wrong fertilizer is an imbalance in the levels of nutrients, both macro and micro. Reduced yield output due to nutritional deficiency leads to an increase in production costs. Because of the deficiency of nutrients in gaseous form, this simultaneously impacts environmental costs. Indian farmers are utilizing soil to its fullest potential, producing two crops annually without using soil management techniques. Over time, the approach alters the soil's chemical composition and causes nutritional deficiencies. These can be taken because the soil loses its crop quality and becomes more prone to microbial contamination. Soil development activities are hindered when there is an imbalance of nitrogen.

In order to keep the soil's organic matter level high and the nutrition abundance abundant, it is currently necessary to conduct continual soil monitoring using a well-understood approach. It is possible to utilize a soil test to determine the nutrient levels, including phosphorus and nitrogen, prior to applying fertilizer. In the event of a nutritional deficit, it is necessary to supply the necessary nutrients. Soil fertility is preserved, and yield output is maximized during this process.

The primary topics this article covers are finding the Soil Grade and Crop Recommendation [12]. The regression algorithm finds the soil grade by considering the various soil nutrients; the Gradient Descent Algorithm minimizes the cost function. This helps farmers comprehend the soil quality. This paper employed many supervised machine learning algorithms, including Random Forest Classifier, Naïve Bayes, and Support Vector Machines, to train a model to recommend the most appropriate fertilizer based on the given soil details. The Random Forest Classifier achieved the highest accuracy rate of 72.74%. Moreover, this study concentrates on specified crops.

Users in South Indian states such as Karnataka, Kerala, Tamil Nadu, and Andhra Pradesh are the target audience for the paper's crop recommendation [13]. Instead of asking for details about soil nutrients and quality, this article asks for information about the area, specifically the state, district, and season. In this study, the model is trained using a random forest classifier. This study includes the development of a website that allows users to get model predictions after registering and logging in. It also provides the crop's ideal rainfall, temperature, and pH. The feasibility of crop rotation is enhanced by offering alternative crops that can be cultivated while considering the expected crop type. This method employs only one machine learning algorithm, and the accuracy point is not considered.

According to the NPK value, inorganic fertilizers are the only ones suggested in [14]. It produces a report on the soil. However, there are a few downsides to this method. It only examines NPK values, has a predefined list of crops, suggests inorganic fertilizers exclusively, and does not have a mobile app. The study by Kiran Shinde, Jerrin Andrei, and Amey Oke in [15] offers advice to farmers on what crops to grow, how to rotate their crops, and which fertilizers to use.

Agriculturalists can access the system through desktop computers, laptops, and smartphones. One disadvantage is that it does not make use of micro level parameters. And it is only available in one language. A lab-on-a-chip system was created to monitor soil nutrients in real-time. Using the concentration of soil nutrients as a calibration, the chip measures changes in charge by capillary electrophoresis. The equipment accurately measured the amounts of NO3, PO4, K, and NH4 ions [16]. This method is cost effective.

Priya and Ramesh (2018) [17] acknowledge that different agricultural locations have distinct soil types. As an additional component of the soil monitoring technique, it has been designed to test the effects of various climatic factors on various crops. For instance, it has been found that some nutrients are more effective in certain climates. Improving productivity and nutrient utilization are two of its most important uses in soil engineering. Global Positioning System (GPS) technology facilitates access to various soil types, sensors, and frameworks for automated hardware control and irrigation systems. GPS-enabled precision farming can provide useful data for better agricultural and environmental management in many regions. Applying water control and spraying drones, as well as the decision support system, benefit from this. This method has proven to be costly for the farmers.

A machine learning-based system that can accurately anticipate crop values and provide recommendations based on such predictions has been suggested in [18]. Aside from that, in the past, Sharma et al. [19] would forecast unpredictable rainfall that impacts crops. In addition, the agricultural yields were predicted by Khan and Ghosh using data from the Meteorological Data of Chhattisgarh (CG) [20]. Various levels of crop nutrients are depicted in the data [20]. This study offers a regression model based on neural networks to forecast when it will rain in the specified region. The information was retrieved from the weather station in Ahmednagar, India. Climate data from the last decade and values, humidity, and precipitation totals are available.

According to Bendre et al. [21], the chosen geo-location rainfall forecast can be enhanced using the regression model. Datasets from the agricultural sector have also been subject to various data mining methods. One example is the clustering-based approach that Hot and Popovic-Bugarin (2015) presented in [22], which incorporates fuzzy k-means. The collected sensor measurements group the soil according to its characteristics. The results were compared with Google Maps and local street segmentation maps in the publication [22], which confirmed that the model was appropriate for presenting data to researchers and farmers. Navarro-Hellín et al. [23] proposed Adaptive Neuro-Fuzzy Inference Systems (ANFIS) for irrigation management in agriculture. The model could foretell the amount of water the irrigation system needs based on the state of the soil, the weather, and the crops. All the above methods discussed do not concentrate on the appropriate fertilizer recommendations.

The drawbacks of fertilizer and pesticides are almost identical. The soil and food crops suffer when insecticides are applied excessively. In [24], the authors check the soil's chemical composition using wireless sensors and then recommend the best time to discharge pesticides based on that reading [25]. Several agricultural research institutes in India worked together to undertake a study that would inform a fertilizer recommendation system that considers soil nutrient content. Use the report to determine the fertilizer needed for each crop in various districts and soil types.

The purpose of this study is to propose a method for crop selection that takes weather and soil conditions into account in order to maximize agricultural production (see [26]). Two primary points are covered. The first is using seasonal weather prediction to determine an appropriate crop. Weather forecasts are based on data collected from the NRSA Hyderabad station, which records meteorological variables such as temperature, humidity, sun hours, wind direction, and more for five years. Using Recurrent Neural Networks (RNNs) for seasonal weather forecasting is not new. The crop dataset includes soil characteristics, weather conditions (such as temperature and humidity), and crop information. Acquired weather forecasts are utilized for crop prediction purposes. Crop and yield predictions are made using machine learning models such as K-Nearest Neighbor, Decision Tree, and Random Forest. Although it is not stated in the study, Random Forest is expected to provide the best results in terms of accuracy. Including fertilizer advice would lengthen the report.

In [27], the authors of this study set out to develop a model that could predict crops given certain inputs; specifically, they offer a web app that could take data like temperature, season, pH, nitrogen, potassium, and phosphorus and feed it into a trained model, which would then provide a crop recommendation as an output. Short descriptions of the climate and soil types ideal for growing the suggested crop are also included on the pages with the crop type itself. This paper only recommends 8 types of crops, and the model is only trained with the KNN. No mention is made of the accuracy that was attained in this study.

In the study [28], as part of their recommendation systems, companies deploy AI and ML algorithms and sensors to detect various parameters. A Machine Learning (ML) model-based crop recommendation system has been created, which utilizes information about the changing seasons, geographical location, and planting season. Compared to conventional farming methods, higher yields from sensors and processors are driving automation in the agricultural sector. Proper insecticides and fertilizers can help farmers achieve optimum output while maintaining healthy crops. A recommendation system should be developed to assist them in this endeavor.

The suggested method 'FertRec' overarching goal is to boost agricultural output and efficiency by providing smart technology to advise farmers on the best fertilizers for crop use. This article is further organized as follows: Section 3 discusses Materials and Proposed Methodology. The implementation part is discussed in Section 4. While experimental results and performance evolution of the proposed system are highlighted in Section 5. The article is concluded in Section 6 with suggested future extensions, followed by references.

3. Material and Methods

Fertilizer recommendation systems have become essential for farmers to enhance their yield. The proposed system 'FertRec' uses two datasets, one for recommending the amount of fertilizers to use and another for recommending the fertilizer type. Later Machine learning models like Decision Tree, SVM Classifier, Logistic Regression, Random Forest, KNN Means and Gradient Boost Algorithms are used for recommending the appropriate fertilizer based on the soil characteristics using the above datasets. The next manual process of recommending the fertilizer based on soil test is explained below to emphasize the importance of the proposed ' FertRec' method.

3.1. Manual Process

A soil test could be any of numerous types of soil analysis performed for various reasons. When it comes to fertilizer recommendations in agriculture, one of the most common reasons for testing the soil is to identify the crop-specific accessible concentration of nutrients. It is possible to conduct soil tests in labs.

Nutrients for crops can be detected in the lab in three main ways [29]:

- Main Nutrients: Nitrogen (N), Phosphorus (P), and Potassium (K).
- Secondary Nutrients: Sulphur, Calcium, and Magnesium.
- Minor Elements: Chlorine, Iron, Manganese, Brass, Zinc, Boron, Molybdenum

The nutritional deficits can induce leaf yellowing or browning, sometimes appearing in unique patterns. Growth retardation and ineffective fruiting or blossoming could accompany this problem. The same is depicted in Figure 1 below. This will adversely affect the farmers and yield. Advising farmers in their native tongue and suggesting the best fertilizers will boost agricultural output.

Fig. 1 Yellowing of leaves due to nutritional deficiency

Currently, farmers are provided with soil test reports through a manual procedure. These reports include information on the soil's available nutrients, recommended fertilizers, and advice in the local language. When done manually, soil test reports might take up to a week to produce. Below is an explanation of the soil tests and fertilizer recommendation system done manually.

The Krishi Vignana Kendra, KVK center's laboratories, tests the soil. Chemical analysis of soil samples reveals the presence of usable nutrients. Table 1 shows nutritional values, whether the available nutrients are minimal, average, or maximum. Based on this nutritional information from the soil test, the farmers consult KVK for recommendations on fertilizers. Land needs, crop types, and fertilizer specifications are all detailed in KVK.

Consider a farmer who wants to cultivate a groundnut crop. They use the following manual procedure to get advice about recommendations of the suitable fertilizers for that crop.

- The specified crop need for fertilizers with the following recommended doses, as suggested by the KVK:

  N=600 gm, P=300gm, and K=300gm
- Nutrient availability determines the dosages.
- As per the given information in Table 1, if the available N in the soil is 189.2 grams, then that inorganic N should be 25% higher than the recommended fertilizer, i.e. 25% more than the specified dose of 625 gm., so

  Recommended N= (600*125)/100=750 gm (1)
- There should be two equal doses of the recommended N, which is 600 times 125 divided by 100, or 750 grams, as stated in the above equation.
- Assuming a soil P availability of 9.94 grams, the inorganic P content should be 125% higher than the recommended fertilizer, as shown in Table 1. Then

  Recommended P= (200*125)/100=250 gm (2)
- Assuming 258 gm of accessible K in the soil, the inorganic K content should be 25% lower than the recommended fertilizer, or 75%, according to Table 1.

  Recommended K= (200*75/) 100=150 gm (3)
- The initial coefficient factor for NPK is computed to determine Uria U, Super Phosphate SSP, and Murate of Potash MoP.

  The coefficient factor is 2.17 since 100 kg of U contains 46% nitrogen N, 100/46 = 2.17 (4)

  The coefficient value is 6.25 because 100 kg of SSP contains 16% Phosphate P,

  100/16=6.25 (5)

  Since 60% of MoP is potassium K, the coefficient factor is 1.67 for 100 kg of MoP, 100/60=1.67 (6)
- Following this, the algorithms determine the suggested Uria U, SSP, and MoP.
- Recommended uria is determined as

  By combining the recommendations in (1) and (4),

  Required Uria, U=2.17 * N=2.17 * 750 = 1627.5
- After that, determine the suggested SSP.

  Recommended SSP= 6.25 * Recommended P = 6.25 * 250 = 1562.5 gm, using Equations (2) and (5)
- Finding the suggested MoP, using Equations (3) and (6)

  Recommended MoP= 1.67 * Recommended K = 1.67 * 20 = 334 gm

Then, the final recommendation of fertilizers by KVK, as per the above calculation, is shown in Table 2.

Table 1. Nutrient range and impact

Nutrient PH	Threshold	<6.5	6.5-7.5	>7.5	
	Action	Acidic	Neutral	Neutral	
EC	Threshold	<1	1-2	2-3	
	Action	Promotes Healthy Crop Development	Negatively Impacts Germination	Negatively Impacts Salt-Sensitive Crop Growth	
Nitrogen N	Threshold	<140	141-560	561-700	>700
	Action	Less	Average	Sufficient	High
Phosphors P	Threshold	<7	8-21	22-35	>35
	Action	Less	Average	Sufficient	High
Potassium m :	Threshold	<100	101-200	201-300	>300
	Action	Less	Average	Sufficient	High
Recommended Fertilizer		50 % More	25 % More	50 % Less	25 % Less

Table 2. Recommendations of fertilizers along with dose

Time of Fertilizer (gms)	Nitrogen	Urea	Phosphorus	SSP	Potassium	MoP
Dose 1	325	813.5	200	1562.5	100	167
Dose 2	325	814	-	-	100	167
Total	750	1627.5	200	1562.5	200	334

3.2. Proposed Model - 'FertRec'

The proposed method in the first phase uses two datasets: one with fertilizer composition required for the crop 'Sugarcane' and another with fertilizer recommendations. The characteristics included in both datasets, including District, Soil Color, Soil pH, temperature, humidity, and rainfall, will be used to train the model.

In the next phase, several machine learning models are used to provide recommendations for the fertilizer by training using two datasets.

The proposed method's process flow follows the steps below, depicted in Figure 2. The steps are as follows:

1) Collection of Datasets
2) Pre-processing (Noise Removal)
3) Feature Extraction
4) Applied Machine Learning Algorithm
5) Recommendation System
6) Recommended Fertilizer

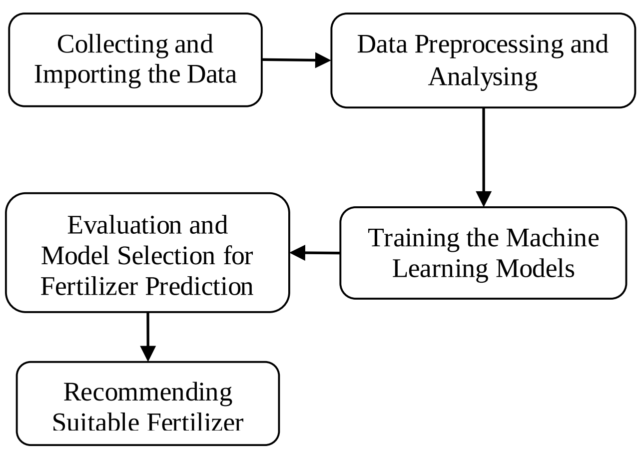


Fig. 2 Process flow of the Proposed method ' FertRec'

The steps are detailed as follows:

1. Two datasets are used to forecast the optimal fertilizer for a specific set of environmental conditions. This dataset includes details about District, Soil color, rainfall, temperature, humidity, pH levels, and fertilizers, among other variables.
2. Datasets are preprocessed after collection to ensure they are acceptable for training the machine learning models. Then, the data is analyzed for any encoding, outliers, and missing or partial data presence.
3. Models are trained with datasets. Accuracy, precision, recall, and F1 score were among the metrics used to assess the performance of the Machine Learning models.
4. The machine learning model shows the highest accuracy and recommends suitable fertilizers for the specified crop.

The process of the proposed method 'FertRec' is depicted in the form of the algorithm below

3.3. Algorithm (Fertilizer Suggestion System)

Procedure_Algorithm (FertRec)

begin
{
//Input 1 : Soil Characteristics: District, Soil Color, Soil pH, temperature, humidity, and rainfall
// Input 2 : Data Sets
//Output : Recommended Fertilizer
Step 1 : Collect the Datasets using the soil characteristics with Soil Test. Dataset 1: Amount of Fertilizer, Dataset 2 : Fertilizer Type
Step 2 : Preprocess the Datasets
Step 3 : Extract the features necessary
Step 4 : Split the datasets for training and Testing with 80% and 20 % weightage.
Step 5 : Train the Machine learning models involved in 'FertRec' Proposed System.
Step 6 : Test the FertRec System
Step 7 : Use the model with high accuracy for fertilizer recommendation.
}
end

4. Implementation

This section uses the datasets originally prepared using different nutrients added to crops over the last decade. Both datasets specify the best and highest possible values. Using the above datasets, machine learning models are later trained to recommend the appropriate fertilizer. All the tests are conducted on a system with 16 GB RAM, Core i7 CPU, google colab, anaconda, and Linux Mint distribution. Maximum

4.1. Dataset Preparation

4.1.1. Dataset – Specifes the Amount of Fertilizer Required

N, P, K values are generally obtained from the soil test. Additional soil factors using a program are computed for data pretreatment and congregated into a dataset. We determine the soil's individual ppm value using the NPK ratio and the total ppm value.

The following formulas are used to find the ppm value for each nutrient derived from the reference [14].

- For Nitrogen: PPM N= 13.1925 * % of N (7)
- For Phosphorus: PPM P=5.8047* % of P (8)
- For Potassium: PPM K=10.949 * % of K (9)

Soil nutrient content and fertilizer application estimates are expressed in kg/HA, the global unit of measurement. Soil nutrient concentration can be estimated by converting ppm to kg/ha. Use to transform the ppm value (obtained in the preceding computation) into kg/HA unit.

Nutrient (kg/HA) = 2.5* PPM of Nutrient (10)

The data was collected by subjecting multiple soil samples from various farming fields to soil analysis. Data is stored in the below format, as shown in Table 3, and may be analyzed and clustered later.

Table 3. Sample soil data analyzed

Label	N	P	K	N (ppm)	P(ppm)	K(ppm)	SN(kg/H A)	SP(kg/H A)	SK(kg/H A)
1	9	10	11	119	58	112	297.5	145	280

SN, SP, and SK are the soil nutrients, nitrogen, potassium, and phosphonate in Kg/Ha. Following data preparation, the soil sample data frame is in tabular format. The fertilizer needed for each crop varies with factors such as soil nutrient level, soil type, region, variety, and season. The dataset also includes the formula for determining the optimal nutrient content of the fertilizer, which is included in the regression.

The following is a sample regression equation:

For Nitrogen content estimation: FN = 4.63 T - 0.56 SN (11)

For Phosphorus content estimation: FP2O5 = 1.98 T - 3.18SP (12)

For Potassium content estimation: FK2O = 2.57 T - 0.42 SK (13)

In this context, T stands for the desired yield, FN, FP2O5, and FK2O denote the necessary fertilizer content in kg/HA, and SN, SP, and SN denote the soil nutrient level in kg/HA.

Using a 9:10:11 soil sample ratio as an example,

The ppm values will be

Ppm N= 13.1925 * 9 = 119
Ppm P= 5.8047 * 10 = 58
Ppm k= 10.949*11 = 112

The nutrient value in kg/HA

SN(kg/HA) = ppm N * 2.5 = 297.5
SP(kg/HA) = ppm P * 2.5 = 145
SK(kg/HA) =ppm K * 2.5 = 280

For the following requirements:

Crop: rice
Soil: Black
State: Telangana
Target yield= 60q/HA

FN = 2.3T – 0.32SN = 2.3 *60 – 0.32 * 297.5
= 138 – 95.2 = 42.7
FP = 1.91T – 1.9SP = 1.91*60 – 1.9*145
= 114.6 – 275 = 0(since negative)
FK = 2.27T – 0.27SK = 2.27*60 – 0.27*280
= 136.2 – 75.6 = 60.

So, finally, the dataset is composed using the above computations as given in Table 4.

Table 4. Sample dataset for the amount of fertilizer recommendation

Soil Label	%N	%P	%K	Required FN	Required FP	Required FK
1	9	10	11	42.7	0	60.5

The above dataset trains the machine learning models discussed in the next section to recommend the required fertilizers.

4.1.2. Dataset - Specifies the Type of Fertilizer

Using a soil test, the dataset for fertilizer recommendation is generated based on the soil characteristics collected in various areas. Utilizing the formula provided in Equation (14), the deficit nutrients and the percentage by which they are deficient are evaluated. This deficiency of the nutrients computed allows us to suggest the appropriate fertilizer in the dataset based on its nutrient composition.

% Deficit for Nutrient for Soil j, DNiSj = (avg(NiCu -NiCl) - SNi)/ avg(NiCu -NiCl) ×100 (14)

Si stands for the jth piece of Soil that belongs to Land L. According to Soil Card data, the real nutrient value, ith property for Soil S, is SNi, NiCu is the upper range of nutrient i for crop C, and NiCl is the lower range.

Deficiency is computed for the soil nutrients for a particular crop, and suitable fertilizer based on the deficiency is incorporated into the dataset.

For Example,

Crop: Rice
Nutrient: Nitrogen

Then, Based on the soil test,

Rice – Nutrient – Nitrogen – Upper value required,
NiCu = 40

Lower Value required,
NiCl = 20

Where, Ni, Nutrient = Nitrogen, C, Crop = Rice, and in Soil, Nutrient, Nitrogen Component, SNi = 10.

% Deficit for Nutrient for Soil from place Pebair, DN (Nitrogen) S (Pebbair) = (Avg(40-20)-10)/(Avg(40-20))*100
= 50 percent of nitrogen is deficient in that soil for cultivating the crop – rice.

Hence, a suitable fertilizer is ‘ Urea’.

After compiling a list of shortfall qualities for Soil S and Crop C, top N fertilizers with a match content percentage of the provided properties are identified and incorporated in the dataset. In this way, the dataset is generated with features that include instance No, Place, Soil Color, Nitrogen(N), Phosphorous(P), Potassium (K), Soil pH, Rainfall, Temperature, Crop, Fertilizer. Sample Dataset is shown in Table 5. For exection I was used Dataset.head().

Table 5. Sample Dataset for recommendation of Fertilizer

Sl. No.	District_name	Soil color	Nitrogen	Phosphorus	Potassium	pH	Rainfall	Temperature	Crop	Fertilizer
0	Pebbair	Black	75	50	100	6.5	1000	20	Sugarcane	Urea
1	Pebbair	Black	80	50	100	6.5	1000	20	Sugarcane	Urea
2	Pebbair	Black	85	50	100	6.5	1000	20	Sugarcane	Urea
3	Pebbair	Black	90	50	100	6.5	1000	20	Sugarcane	Urea
4	Pebbair	Black	95	50	100	6.5	1000	20	Sugarcane	Urea

The process of preparing the dataset to identify the appropriate fertilizer is given in the form of the algorithm below.

```
Algorithm (Dataset_Fertilizer Suggestion)
Procedure_Algorithm (Dataset_FertRec)
begin
{
//Input : Soil Nutrients, Location, Crop, CropNutrients _Upper, CropNutrients_Lower
//Output : Suggested_Fertilizer, Dataset
// Real nutrient value of Soil S - > SNi
// Upper range of nutrient i for crop C - > NiCu
// Lower range of nutrient i for crop C - > NiCl
// % Deficit for Nutrient for Soil j = DNiSj
If(SNi < NiCl), then
  Compute
  Absolute Deficit for Nutrient for Soil j, ADNiSj=
              (avg(NiCu -NiCl ) - SNi)
  % Deficit for Nutrient for Soil j, DNiSj = ADNiSj /
   avg(NiCu -NiCl) ×100
Recommended_Fertilizer, RF=
NearestComposition_Fertilizer(% Deficit DniSj,  F)
}
end
```

4.2. Machine Learning Models

Several Machine Learning methods, including Decision Tree, Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), KNN means, and Gradient Boost, were utilized in this proposed system. The two datasets prepared above are used to train and test the models mentioned above as part of the proposed FertRec method.

The preprocessed data is normalized to apply Machine Learning algorithms. Once the data training is finished, the model may be tested on the test dataset to see if the prediction is correct. Compared to the accuracy levels of the individual algorithms, the one with the relatively high accuracy will be considered for fertilizer recommendation. This way, based on the user's input datasets, the proposed 'FertRec' model recommends fertilizers for the crop.

After the model is trained using ML methods, it can forecast values at runtime based on user-supplied new values. Machine learning projects are hosted on the stream light framework, which is used to implement the complete project. Below are the details of the machine learning algorithms used in the proposed model.

4.2.1. Decision Tree

The dataset was trained using a Decision Tree method. The decision tree approach divides the data set containing the desired attributes into progressively smaller nodes using a supervised learning algorithm. Each node in a tree has three roles: root, decision, and terminal. The variance error source, which increases as the model complexity increases, can be reduced using tree pruning and time series cross-validation. In order to discover the best recursive binary node splits, this technique uses a greedy top-down strategy that minimizes variance at the terminal node locally. Classifiers that use decision trees often employ greedy algorithms. Using a tree to encode attributes and class labels, it is a supervised learning algorithm. Using decision rules inferred from training data, the primary goal of a Decision Tree is to build a training prototype that can predict the class or value of target variables.

4.2.2. SVM

The best crop to grow can be predicted using the SVM method for classification, which sorts the many soil factors. In order to assess the soil properties and suggest an appropriate crop, the suggested algorithm is run in anaconda navigator. When it comes to classification, the SVM method is being evaluated.

4.2.3. Logistic Regression

One popular statistical model is the Logistic Regression model, which has many more advanced variants, but at its core, it models a binary dependent variable using a logistic function. Logistic regression, a subset of binomial regression, is used to forecast logistic model parameters in regression analysis.

4.2.4. Random Forest

A simple and adaptable approach that yields strong predictions is Random Forest regression. One prediction-making machine learning technique is random forest regression.

4.2.5. K Nearest Neighbor (KNN)

Crop recommendations are made using the K Nearest Neighbor (KNN) algorithm. Input characteristics such as soil type, land type, soil texture, nitrogen, phosphorus, and potassium are first collected from the user. After that, we will filter the dataset according to the user-supplied soil type, land type, and soil texture.

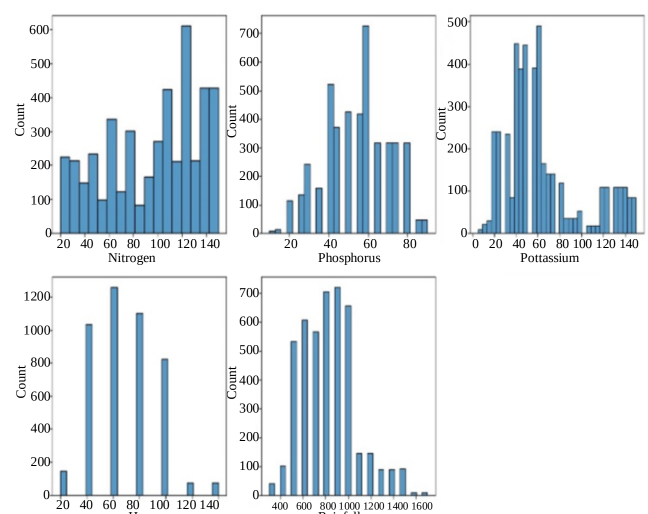
4.2.6. Gradient Boost

In contrast, XGBoost is a gradient boosting method that constructs a group of weak decision trees and then uses their combined predictions for a more precise outcome.

5. Experimental Results and Performance Analysis

The CSV-formatted Fertilizer datasets have been cleaned and prepared for data frame training. In this case, the training dataset is 80% larger than the test dataset, which is 20% smaller. Both datasets include 100 tuples with all the mentioned soil characteristics. Machine learning models are applied to the datasets to recommend the appropriate fertilizer based on soil features. Results are presented and discussed in this section. Features in Datasets are visualized in the Figure 3.

We use the confusion matrix to find out how well the categorization models did on a certain set of test data. If the test data's actual values are known, they can be calculated. The heat map is given in Table 6. Table 7 represents The Precision, recall and F1 score values of the various machine learning algorithms.



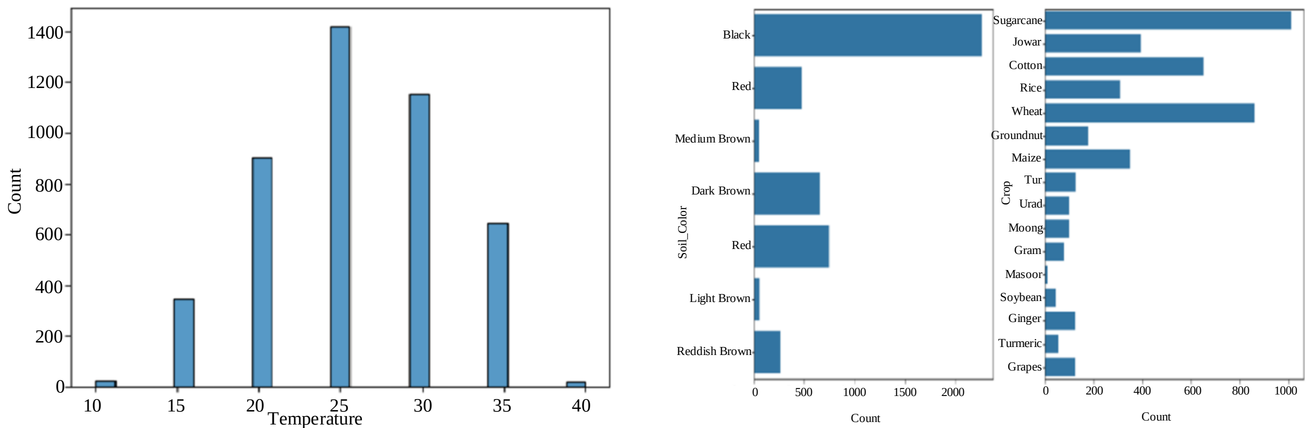


Fig. 3 Visualization of features in dataset: nitrogen, phosphorus, potassium, soil color, crop, temperature, rainfall, soil ph values

Table 6. Heat map generated for the proposed model

	Nitrogen	Phosphorus	Potassium	pH	Rainfall	Temperature
Nitrogen	1.000000	0.700539	0.584315	0.182850	0.269364	-0.010213
Phosphorus	0.709539	1.000000	0.573970	0.244945	0.225453	-0.055303
Potassium	0.584315	0.573970	1.000000	0.075110	0.445671	0.053413
pH	0.182850	0.244945	0.075110	1.000000	0.097884	-0.002949
Rainfall	0.269364	0.225453	0.445671	0.097884	1.000000	0.315045
Temperature	-0.010213	-0.055303	0.053413	-0.002949	0.315045	1.000000

Table 7. The Precision, recall and F1 score values of the various machine learning algorithms

	Precision	Recall	F1 Score
Decision Tree	1	1	1
Logistic Regression	1	1	1
Random Forest	1	1	1
KNN	0.87	0.9	0.88
Gradient Boost	0.97	0.95	0.96

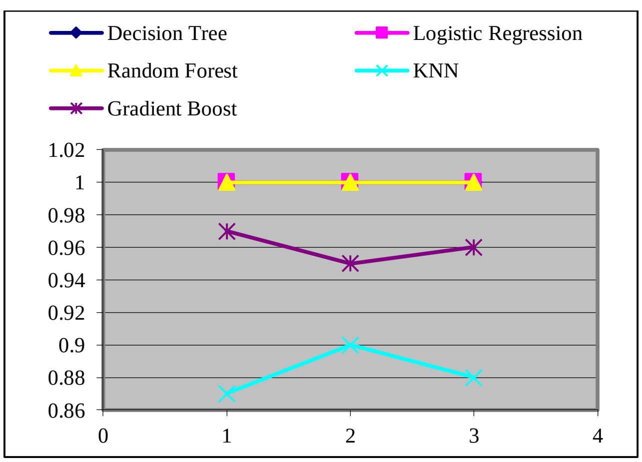


Fig. 4 The precision, recall, and 1 score values of the various machine learning algorithms

The accuracy comparison of the applied machine learning algorithms is shown in Figure 5 below. And the same is represented in Table 8.

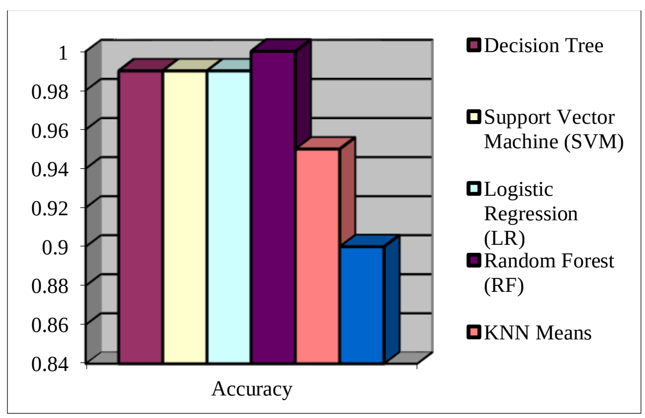


Fig. 5 Accuracy comparison of all algorithms

Table 8. Accuracy stats of various machine learning algorithms

S. No.	Machine Learning Model	Accuracy
1	Decision Tree	0.99
2	Support Vector Machine (SVM)	0.99
3	Logistic Regression (LR)	0.99
4	Random Forest (RF)	1
5	KNN Means	0.95
6	Gradient Boost	0.9

From the above results and accuracy table, it is observed that the Random Forest model has shown better accuracy compared to other algorithms. So, Random Forest is chosen in the proposed fertilizer recommendation model, 'FertRec,' to recommend suitable fertilizers based on the datasets and soil features supplied. The fertilizer recommendations of the Random Forest Machine Learning Model are depicted in Table 9.

Table 9. Recommendation of fertilizers – random forest algorithm

Label	District	Soil Color	N	P	K	pH	Rainfall	T	Crop	F
1	Pebbair	Black	90	25	25	6.0	1500	35	Sugarcane	DAP
2	NGKL		140	40	50	6.0	1600	25	Sugarcane	Urea
3	MBNR	Brown	115	60	45	7.0	800	25	SC	19:19:19 NPK
4	GDWL	Black	60	80	145	5.0	1000	35	SC	Urea

The performance analysis of the proposed fertilizer recommendation system, "FertRec", is evaluated by comparing it with the fertilizer recommendation system without machine learning algorithms, FRS proposed in [12], in terms of access time and sample size. The efficacy of the proposed system, FertRec, is shown in the graph below, in Figure 6, and the stats are depicted in Table 10.

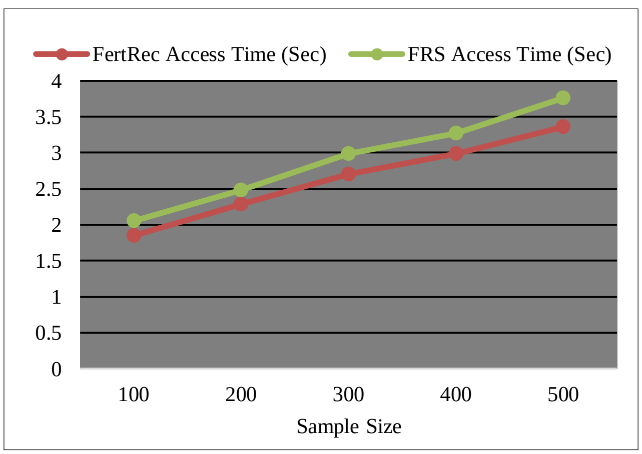


Fig. 6 Performance analysis of the FertRect – Access time

Table 10. Access time in Sec FertRec vs FRS

Sample Size	FertRec Access Time (Sec)	FRS Access Time (Sec)
100	1.85	2.05
200	2.28	2.48
300	2.7	2.987
400	2.98	3.265
500	3.36	3.756

The results of the FertRec proposed model reveal that the processing time for fertilizer recommendation tends to grow as the data size increases (Figure 6). However, it outperforms the existing fertilization recommendation model FRS. The time needed to suggest fertilizers has decreased when the proposed 'FertRec' method is implemented based on machine learning. When dealing with larger data sets, improving algorithm performance and increasing processing time are common outcomes. Regardless, in comparison to more conventional methods of fertilizer recommendation, the proposed FertRec method outperforms the competition.

6. Conclusion

Nutrient management and fertilizer recommendation systems are addressed in this article, which uses two data types and several machine learning algorithms. The most recent and cutting-edge machine learning methods for fertilizer recommendation and datasets are the main subjects of this article. It talks about how machine learning could help with fertilizer recommendation prediction. Soil samples taken from agricultural land can have their NPK ratios calculated using the suggested approach, which then suggests the optimal fertilizer to use in the right amounts to nourish the soil for the chosen crop. Without harming the land or soil qualities, the suggested system aids farmers in getting the most out of each crop cycle. Because the likelihood of over-fertilization is reduced, this also guarantees that healthy crops have been grown.

This research shows that the suggested approach, FertRec, works better and yields more fruits than the manual process and other methods. It turns out that the suggested model can get you a greater nutritional quantity. Fertilizing crops with the correct amounts of nutrients allows for a good yield. The suggested model brought about increased crop yield output and enhanced capacity to choose optimal combinations of available resources. This will be useful for agricultural professionals and farmers who want to use the same approach with different crops. To save time and effort, the N-P-K has been mechanized. The technique works better if a large dataset of various crops is available, and it aims to achieve yield figures that are appropriately indexed with soil type and location.

The proposed FertRec Fertilizer recommendation system can analyze unique soil types and provide tailored recommendations. In addition, compared to conventional fertilizer recommendation techniques, the algorithm's execution time is four times faster on average. As a result, we conclude that this system outperforms other benchmark algorithms when it comes to fertilizer recommendation. Although machine learning has proven successful in fertilizer recommendation systems, there is a lack of research on identifying illnesses in recommended crops, providing pest control approaches for forecast crops, and recommending fertilizers. So, this method can be extended to predict when pests will be a problem shortly.

Acknowledgement

The authors acknowledge the direct and indirect support and cooperation rendered by all the members.

References

[1] Alejandro Castañeda-Miranda, and Victor M. Castaño-Meneses, "Internet of Things for Smart Farming and Frost Intelligent Control in Greenhouses," *Computers and Electronics in Agriculture*, vol. 176, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Yakindra Prasad Timilsena et al., "Enhanced Efficiency Fertilisers: A Review of Formulation and Nutrient Release Patterns," *Journal of the Science of Food and Agriculture*, vol. 95, no. 6, pp. 1131-1142, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[3] Poreddy Ishika Reddy et al., "Automated Plant Disease Detection Using Convolutional Neural Networks: Enhancing Accuracy and Scalability for Sustainable Agriculture," *International Journal of Computer Engineering in Research Trends*, vol. 11, no. 9, pp. 1-10, 2024. [Publisher Link]

[4] Xiaohui Chen et al., "What has Caused the Use of Fertilizers to Skyrocket in China?," *Nutrient Cycling in Agroecosystems*, vol. 110, pp. 241-255, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[5] Huang Shao-Wen et al., "Reducing Potential of Chemical Fertilizers and Scientific Fertilization Countermeasure in Vegetable Production in China," *Journal of Plant Nutrition and Fertilizers*, vol. 23, no. 6, pp. 1480-1493, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[6] Yuanzhi Guo, and Jieyong Wang, "Spatiotemporal Changes of Chemical Fertilizer Application and Its Environmental Risks in China from 2000 to 2019," *International Journal of Environmental Research and Public Health*, vol. 18, no. 22, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Walter Ocimati, Sivalingam Elayabalan, and Nancy Safari, "Leveraging Deep Learning for Early and Accurate Pre-diction of Banana Crop Diseases: A Classification and Risk Assessment Framework," *International Journal of Computer Engineering in Research Trends*, vol. 11, no. 4, pp. 46-57, 2024. [Publisher Link]

[8] Xiang-de Yang et al., "Effects of Long-Term Nitrogen Application on Soil Acidification and Solution Chemistry of a Tea Plantation in China," *Agriculture, Ecosystems & Environment*, vol. 252, pp. 74-82, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[9] Tang Han et al., "Research Progress Analysis on Key Technology of Chemical Fertilizer Reduction and Efficiency Increase," *Transactions of the Chinese Society of Agricultural Machinery*, vol. 50, no. 4, 2019. [Google Scholar] [Publisher Link]

[10] Xiaoying Yang, and Shubo Fang, "Practices, Perceptions, and Implications of Fertilizer Use in East-Central China," *Ambio*, vol. 44, pp. 647-652, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[11] Wenchao Li et al., "Comprehensive Environmental Impacts of Fertilizer Application vary among Different Crops: Implications for the Adjustment of Agricultural Structure Aimed to Reduce Fertilizer Use," *Agricultural Water Management*, vol. 210, pp. 1-10, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[12] Krupa Patel, and Hiren B. Patel, "Multi-Criteria Agriculture Recommendation System Using Machine Learning for Crop and Fertilizesrs Prediction," *Current Agriculture Research Journal*, vol. 11, no. 1, pp. 137-149, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Chaitanya B.N. et al., "Features Identification for Growth of Certain Crops in Indian Agriculture," *International Journal of Creative Research Thoughts*, vol. 9, no. 8, pp. 735-739, 2021. [Publisher Link]

[14] Online Fertilizer Recommendations. [Online]. Available: http://stcr.gov.in/Farmer/main.aspx

[15] Kiran Shinde, Jerrin Andrei, and Amey Oke, "Web Based Recommendation System for Farmers," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 3, no. 3, pp. 1444-1448, 2015. [Publisher Link]

[16] R.G.V. Bramley, and J. Ouzman, "Farmer Attitudes to the Use of Sensors and Automation in Fertilizer Decision-Making: Nitrogen Fertilization in the Australian Grains Sector," *Precision Agriculture*, vol. 20, pp. 157-175, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[17] Rashmi Priya, Dharavath Ramesh, and Ekaansh Khosla, "Crop Prediction on the Region Belts of India: A Naïve Bayes MapReduce Precision Agricultural Model," *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, India, pp. 99-104, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[18] G. Suresh et al., "Efficient Crop Yield Recommendation System Using Machine Learning for Digital Farming," *International Journal of Modern Agriculture*, vol. 10, no. 1, pp. 906-914, 2021. [Google Scholar] [Publisher Link]

[19] Amit Kumar Sharma, Sandeep Chaurasia, and Devesh Kumar Srivastava, "Supervised Rainfall Learning Model Using Machine Learning Algorithms," *The International Conference on Advanced Machine Learning Technologies and Applications*, pp. 275-283, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[20] Huma Khan, and S.M. Ghosh, "Crop Yield Prediction from Meteorological Data Using Efficient Machine Learning Model," *Proceedings of International Conference on Wireless Communication*, pp. 565-574, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[21] R.S. Loomis, J. Rockström, and M. Bhavsingh, "Synergistic Approaches in Aquatic and Agricultural Modeling for Sustainable Farming," *Synthesis: A Multidisciplinary Research Journal*, vol. 1, no. 1, pp. 32-41, 2023. [Google Scholar] [Publisher Link]

[22] Elma Hot, and Vesna Popović-Bugarin, "Soil Data Clustering by Using K-Means and Fuzzy K-Means Algorithm," *2015 23rd Telecommunications Forum Telfor (TELFOR)*, Belgrade, Serbia, pp. 890-893, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[23] H. Navarro-Hellín et al., "A Decision Support System for Managing Irrigation in Agriculture," *Computers and Electronics in Agriculture*, vol. 124, pp. 121-131, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[24] Kasula Kedhari Priya et al., "Towards a Greener Tomorrow: The Role of Data Science in Shaping Sustainable Farming Practices," *International Journal of Computer Engineering in Research Trends*, vol. 11, no. 4, pp. 12-19, 2024. [Publisher Link]

[25] R.S. Dinesh, and B. Ramamoorthy, "*Soil Test Crop Response Based Online Fertiliser Recommendations*," IISS, Bhopal, 1968. [Publisher Link]

[26] Anu Bala, "Machine Learning Approaches for Crop Yield Prediction-Review," *International Journal of Computer Engineering and Technology*, vol. 11, no. 1, pp. 23-27, 2020. [Google Scholar] [Publisher Link]

[27] U.M. Prakash et al., "KNN-Based Crop and Fertilizer Prediction," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 4, pp. 1453-1456, 2020. [CrossRef] [Publisher Link]

[28] M. Bhavsingh, Y. Alotaibi, and S. Alghamdi, "Fusion of Convolutional Neural Networks and Gradient Boosting Machines for Spinach Leaf Classification and Prediction," *International Journal of Computer Engineering in Research Trends*, vol. 11, no. 3, pp. 38-45, 2024. [Publisher Link]

[29] Tanmay Thorat, B.K. Patle, and Sunil Kumar Kashyap, "Intelligent Insecticide and Fertilizer Recommendation System Based on TPF-CNN for Smart Farming," *Smart Agricultural Technology*, vol. 3, 2023. [CrossRef] [Google Scholar] [Publisher Link]