Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

Artificial Intelligence and Machine Learning based Farmer's Friendly Soil Fertilizer Recommendation System Through Expert Analysis

[1]Anil Kumar Ghadiyaram, *[2]Dr. Srinivasa L Chakravarthy

[1]Assistant professor, Department of CSE, Vignana Bharathi Institute of Technology, Hyderabad, Telangana, India
[2]Assistant Professor, Department of Computer Science and Engineering, GITAM (Deemed to be University), Visakhapatnam, AP, India
[1]anilkumarghadiyaram@gmail.com, [2]chakri.ls@gmail.com

Article History:

Received: *26-09-2024*

Revised: *07-11-2024*

Accepted: *26-11-2024*

Abstract:

One of the key components of the Indian monetary system is agriculture and as is well known, farming supports about 60% of the Indian population and accounts for a sizable portion of the country's GDP. As time goes on, India is becoming one of the world's most significant food exporters. The demand for meals has increased as a result of the growing population. Many educated individuals who work as farmers now also work in agriculture. A promising strategy for increasing crop productivity and maximizing use is precision farming which utilizes machine learning (ML) and Internet of Things (IoT). However, due to a decline in productivity, farmers are still facing issues with significantly lower incomes. Even farmers are making poor crop, fertilizer, and soil choices. This study offers a novel solution to this problem by employing professional analysis and ML and artificial intelligence (AI) techniques to create a farmer-friendly soil fertilizer recommendation system. After analyzing the records, the AI machine uses this suggested technique to provide farmers with answers based on expert review and a thorough historical record set. Therefore, farmers may find the crop suggestion systems to be highly advantageous. Additionally, a number of factors, including pH, nitrogen, phosphorus, potassium, and rainfall, can affect crop output. Therefore, we are offering a method in this paper for using AI and ML algorithms to increase crop yield.

Keywords: Artificial Intelligence, Machine Learning, soil analysis, Crop Yield, Food Security, Agriculture, Fertiliser Recommendation.

Introduction

One of the most important sectors of the Indian financial system is agriculture. Regarding employment, more than 151 million people rely on the agriculture zone for their livelihood. Approximately 60% of the Indian population is employed by the company, which accounts for 18% of the country's GDP. Globalization has led to advancements in India's agricultural sector [1]. Farmers have historically learned from other farmers' research and input, yet they frequently struggle to grow vegetation effectively, which can result in a lack of revenue. Nowadays, the world needs to produce high-quality vegetables as people become more fitness conscious. To achieve maximum yield, farmers apply fertilizer and insecticides to the soil.

Pesticides, which include herbicides, fungicides, insecticides, and so on, are toxic substances used to eradicate pests, weeds, fungi, and other organisms. [2]. In addition to being more toxic than herbicides and fungicides, pesticides are chemicals that stop pest attacks on crops. [3]. Under "The insecticide act, 1968," which includes regulations for the safe and appropriate use of pesticides, the manufacture, sale, and import of pesticides are governed. Some farmers are ignorant of the negative effects of oversprinkling insecticides on plants and adding fertilizer to the soil without first determining whether the soil is suitable. Spraying too many pesticides can harm both people and the environment.

Agriculture, having a significant impact on a rural community's economy is deteriorating nowadays as a result of the shifting natural factors. The environmental factors that agriculture depends on instantly include sunlight, humidity, soil type, rainfall, maximum and minimum temperatures, climate, fertilizers, pesticides, and so forth. For agriculture to flourish, knowledge on how to properly harvest vegetation is necessary. Seasons in India include:

1. Winter, which lasts from December to March

2. April to June is summer.

3. Rainy season or monsoon, which lasts from July to September

4. The fall or post-monsoon season, which lasts from October to November.

It's critical to assess whether plants are suitable for domestication due to seasonal and rainfall variations. The main problems that farmers deal with include crop control, crop yield forecasts, and vegetation productivity. Farmers and other cultivators seek the right help with crop production because a lot of young people are interested in agriculture these days. We must change the traditional farming method in order to meet farmers' needs for food and useful items. This method does not accurately determine the soil homes, the amount of water required for a particular crop, the amount of profit that can be made from the crops, or which crops and fertilizers are suitable for the land in question.

Generation is comparable to an advancement in agricultural research and development. Agricultural studies benefit greatly from a variety of technologies, such as digital image processing, device learning, deep learning, and large records. These days, there are many other types of statistics gathering and evaluation methods available, including ML algorithms, recommendation systems, and digital image processing. Additionally, the computational age of ML creates a version that can analyze itself and use the model to make predictions. RC may be made more accurate and efficient with device learning. By suggesting suitable soil, crops, and fertilizers, those technologies could enhance the expansion of the agricultural area. Few researches on the use of era in agriculture were conducted prior to the 20th century.

After the introduction of big data and ML, numerous academics have begun to work in agriculture to address contemporary issues. Cutting-edge technology will help with the efficient manipulation of large amounts of data and processing of that data in order to improve crop production, identify and specify plants, and suggest fertilizers. Three key problems with conventional agriculture have been identified based on our observations.

➢ The challenge of choosing appropriate plants that optimize output and revenue
➢ The issue with fertilizer selection is that it is mostly dependent on crop and soil conditions, with precise ratios

- A challenge in identifying agricultural illnesses and prescribing remedies for them.

A recommender system's goal is to give the user relevant recommendations based on their capabilities. It significantly cuts down on the amount of time the user has to search for the things that most interest them and to find stuff that they might not have noticed but are likely to enjoy. There are many different ways to define recommendation systems. Robin Burke [4] has the most widely used and up-to-date definition, which we apply here: "a recommender device is a system able to supply personalised tips or of guiding the user to interesting or useful sources inside a big information area". Numerous scholars are working on issues pertaining to common agricultural structures and creating algorithms. Selecting the best crop for a peculiar land will be solely based on the number of variables like soil characteristics, climate, water level, land proximity, market demand, and more.

Using an apriori algorithm, farmers can receive plant recommendations based on market assessment. A logistic regression strategy is suggested for each other call-based set of rules. increasing crop productivity with a variety of classifiers, including logistic regression, SVM, and decision trees. In our analysis, we found that the majority of researchers focus on soil houses and brands to suggest plants, and we are creating a system that will have pleasant conversations with farmers. But if you want to boost agricultural productivity, you also need to consider other factors when predicting crop yields for a specific land, such as soil type, soil habitats, water and temperature requirements, land region, and crop market value. number of synonyms suggest CR systems.

We don't overlook words that update "advice" with the help of "choice" or "inspiration" on this gadget. The phrases "AI" and "ML" were used to find papers that employ novel approaches based on agricultural data. The seek question (SQ), which is provided as follows, contained the majority of these phrases:

Query = ((“Crop” OR “Soil” OR “Fertilizer”) AND ("AI" OR "ML") AND (Recommendation OR Suggestion))

Related Work

ML has advanced significantly in many facets of life, and agriculture is no different. With the help of its practical recommendations for resource efficiency and agricultural output enhancement, it has completely transformed farming. As a result, we are combining ML and AI in our research to develop farmer-friendly agricultural guidance devices.

Our system's goals are to increase agricultural yield and quality while using less resources and raising farmer incomes. We employed solely ML techniques and an advanced crop advising system in our earlier research. In the literature, numerous works were suggested for categorizing soil fertility and forecasting the plants that will be consumed by those plants. ANNs, NB, DT, RF, SVM, AdaBoost, NN, and LR are among the system-understanding algorithms used in the majority of the tactics. These algorithms include Bayesian networks with ensemble standards of mixing effects from greater fashions along with boosting, stacking, bagging, and greater for soil type and crop prediction [18], [19].

Table 1: Analysis of Existing Crop Recommendation Models

Model	Methods Used	Advantage	Disadvantage
Classification of Soil depending upon the crop relevancy [2]	Random forest classifier	Helps in suitable crops identification	Doesn’t considered the other environmental factors
Crop prediction model [5]	Recursive feature elimination technique with the adaptive bagging	Considers soil and environmental factors	Doesn’t considered the crop predecessor
Crop rotation and yield analysis [16]	Naïve ratio algorithm	Predicts suitable crops than naïve Bayes	There is no detailed study found on rotation
Crop prediction model [17]	NB classifier	Utilizes the environmental factors	Not given any detailed analysis on the result
Soil fertilizer recommendation system - for paddy fields [19]	GA and PSO-based parameter optimization for an enhanced SVM classifier	Improved accuracy among other variations of SVM	No comparison of results with other classifiers
Model for predicting pH levels and soil fertility indicators [20].	Extreme learning machine	Improved classification result	Results are not compared with other classifiers
Wheat yield potential prediction model [26]	Supervised self-organizing maps	provides better results	Doesn't include climate when forecasting yield.
Crop recommendation system [27]	K nearest neighbour classifier	Displays soil deficiency	Only the properties of the soil are used to predict crops.
Predicting crop suitability [29]	Rough set on fuzzy approximation space	Better classification accuracy	Doesn't took crop predecessors into consideration.
Precision crop suggestion [30]	Deep neural network	Insist on good prediction accuracy	No comparison with other classifier models
System for crop recommendations based on environmental factors [31]	Regression-based ensemble	Most suitable for primary crops	No comparison with other classifier models

Mechanism for recommending crops based on environmental factors [32]	Ensemble classifier	Offers more accurate predictions than XGBoost & random forest.	Lacks in comparison with other classifiers

Crop disease prediction, crop yield prediction, climate forecasting, smart irrigation systems, and figuring out the minimum help price are just a few of the applications of ML models in agriculture that have been highlighted by (Al-Gaadi et al., 2016; Nandy and Singh, 2020; Sharma et al., 2020; Cravero and Sepulveda, 2021). In an effort to make precise predictions, researchers used supervised ML algorithms to forecast crop yield (Kaur, 2016; Shehadeh *et al*., 2021). The DT classifier has been used to produce predictions for agricultural temperature and yield (Lee and Moon, 2014; Bagis *et al*., 2012). KNN and ID3, a DT version, were used to analyze the agricultural productivity over the preceding 12 months (Charbuty and Abdulazeez, 2021).

Several studies employ statistical models, such as the auto-regressive integrated moving average (ARIMA) and ML version SVM, to estimate agricultural productivity (Sujjaviriyasup and Pitiruek, 2013). However, time series evaluation has been used to forecast the rate and production of vegetables and crops. The objective was to investigate supply and demand factors and become aware of a time series feature that may identify trends and seasonality in exact greens.

It is regarded as a new area in this business because of the type of research that has been done on the issue of applying ML to agriculture. Diverse ideologies have been developed and examined in the realms of agriculture and related subjects by researchers from all over the world. [3] SVM was suggested by N. Gandhi, L. J. Armstrong, O. Petkar, and A. ok. Tripathy as a method of forecasting rice crop yield. A variety of variables, including location, temperature, precipitation, and production, are included in the dataset used for this method. Sequential minimal optimization is the classifier employed in this dataset. They arranged the dataset using the Weka tool in order to provide the set of rules on the dominant dataset. Python has evolved the impacts of the SVM algorithm.

[4] S. Veenadhari, B. Misra, and C. Singh created the Crop Recommender website, an interactive tool for evaluating how crop productivity is impacted by climate change. The C4.5 algorithm served as the sole basis for the creation of decision timber and regulations. It demonstrates the wide range of climate factors that impact crop development. Temperature, precipitation, and other environmental factors from the pertinent years were examined. The alternatives were decided by the zones beneath the selected crop.

Rose *et al*.'s study [26] emphasized the importance of ML classifiers and statistical techniques in predicting soil fertility and managing ecosystems with less human involvement. In [27], Rajamanickam used KNN, SVM, and selection trees to forecast soil fertility based only on information about macro- and micronutrients. The accuracy of the decision tree approach was 99%.

An agricultural guidance device was proposed by Garanayak *et al.* [15] using regression-based techniques. One of the main goals of the paper was to forecast agricultural productivity. The following are the algorithms utilized in this paper: linear, polynomial, SVM, DT, and RF regression. Five crops, rice, gram, ragi, onion, and potato were utilized to construct the dataset. The chosen climate factors that influenced crop yield were season, cloud cover, area, and vapor pressure. The majority vote was heavily utilized by the authors to determine what was both enticing and healthful. In a crop-based manner, plots representing area instead of production were also employed for several chosen ML approaches. 94.78% of the votes cast indicated that this article was correct.

Problem Statement

Unfortunately, no publication to yet has warned against farmers using such modern crop recommendation (CR) tactics in a constructive way. The available records are a potential capability hurdle that researchers may encounter while attempting to resolve this delay. For the purpose of managing agricultural risk and forecasting the future, accurate records regarding crop output history are essential. One of the most important agricultural problems is predicting crop yield. Most of the time, climate factors (such as temperature and rainfall) and pesticides affect agricultural productivity. Some recent graduates or college students are now drawn to pictures in agricultural methods that they are unfamiliar with in that area. In order to rely on the crop and soil factors largely based on the previous model, we gathered the beyond information and used the ML set of policies. We must use Farmer's nice CR system with ML and AI as the solution. This could benefit farmers by increasing field productivity and lowering fertilizer use in crop production.

Contributions of the Study

The principal objective of this research is to help farmers choose which crops to grow and how much fertilizer to use. Taking into consideration several relevant factors, the study develops a comprehensive model that integrates CFRS holistically to give farmers more specific and tailored recommendations. Highlighting the study's noteworthy contributions are the following:

Data Processing and Profiling: Comprehensive information about Rwanda's main crops was gathered & examined, with an emphasis on data accuracy and based content material**.**

Correlation Evaluation: The goal of the study is to improve decision-making by investigating how inter-variable correlation can improve predictive modeling.

Crop Recommendation: To recommend crops, the study uses a neural community model. This version is more successful than some well-known ML styles and has been carefully trained and studied**.**

Soil and Fertilizer recommendation: The basic knowledge that every soil and crop aggregation has different nutrient requirements is supported by the modeling of fertilizer advice, which uses a straightforward logical characteristic. Adoption of the established principles, which were based on actual agricultural practices, helped farmers understand the rationale behind recommendations, promoting consideration and adoption.

The study also offers a conceptual framework for putting the proposed CFRS into practice, which includes monitoring soil conditions and nutrient dynamics throughout time, in order to provide practical, fact-based, and real-time agricultural suggestions.

Outline of the Study

The final section of the suggested manuscript is organized as follows: Section 1 provides a brief overview of the related work, indicates the current state of research, and differentiates proposed work from existing work. Section 2 explains how each module of the proposed system uses a soil fertilizer machine that is easy for farmers to use. The performance measures used in the experimental procedure, the evaluation of the final product, and the overall performance are then covered in section 3. The use case scenario of the suggested system with regard to actual deployment scenarios is also provided in this section. Section 4 ultimately brings the job to a close.

Therefore, it may be evident that a great deal of research has been done on PA packages and predictive modeling that is mostly ML-based in order to benefit agricultural systems. Although each approach has its own advantages, it also has limitations. Additionally, it has been determined that, within the framework of Rwanda's agricultural system, relatively little work is completed. There is a wealth of research on various ML applications in smart agriculture in the literature. However, rather than being implemented, such processes are now the focus of theoretical discussion. Although the theoretical debates are useful, it is crucial to validate those models in real-world international agricultural situations for evaluating their viability & efficacy. Furthermore, it has been noted that the current research is devoid of information regarding the implementations of statistics source systems, even if they no longer outline the functions that were employed or the criteria by which they selected a specific mastery model. The suggested machine covered in the following section fills in all of these gaps.

Implementation Procedure

Dataset Preparation

The information was gathered from [16]. Nearly 6000 rows, or samples, and 15 columns, or 15 attributes, make up the dataset. There are 17,600 cases in total. There are seven predictors, and the name of the anticipated crop is included in the response column. The following predictors make up the dataset, namely:

- Phosphorus content of the soil.
- Nitrogen content of the soil.
- Potassium content of the soil.
- pH of the soil.
- Soil Temperature.
- Average rainfall in millimetres.
- Humidity content of the soil.

A farmer-friendly RC prediction model that will recommend the best crops to produce on a particular farm based on a number of parameters was developed using the dataset collected from kaggle.com. The dataset's dimensions are (6000, 15).

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

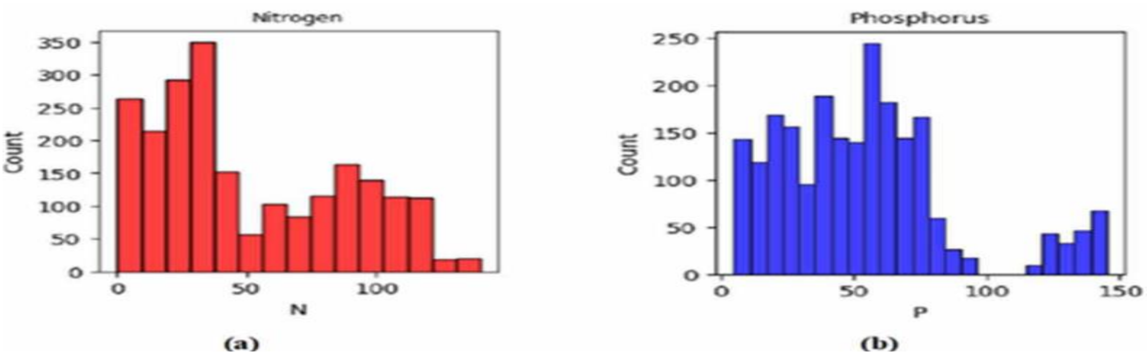

Figure: Nitrogen and Phosphorus for attributes analysis

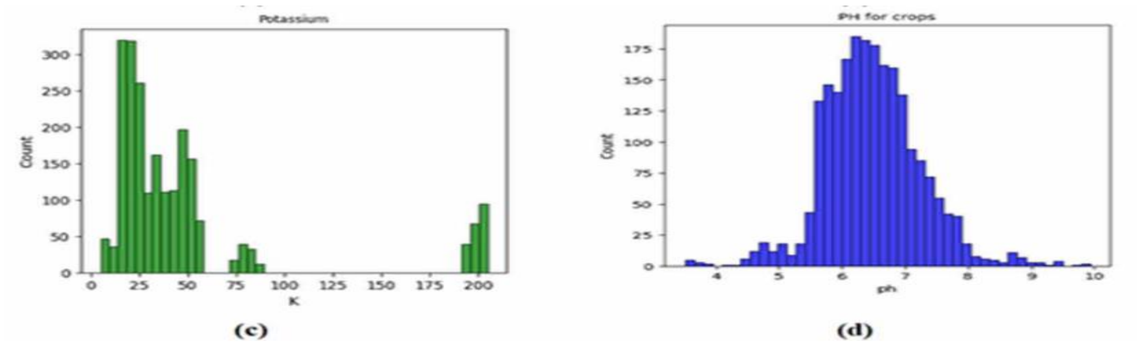

Figure: Potassium and pH for crops attributes Analysis

Using supervised learning, the ML model for crop recommendations was developed. We first pre-processed the data to deal with outliers and missing values, as well as to normalize the features. About 6000 samples from the dataset were separated into training and testing sets with an 80:20 split ratio. Eighty percent of the data were kept for training, and the remaining twenty percent were kept for testing the trained model.

Data Attributes

Agronomic research has demonstrated that a wide range of environmental and geographic factors, such as soil type, humidity, temperature, rainfall, altitude, and more, significantly affect crop development, yield, and fitness. I believe that those factors, either separately or in combination, contribute to determining if a particular crop is appropriate for a particular region. The study has taken into account the developing understanding of the fundamental soil nutrients N (nitrogen), P (phosphorus), OK (potassium), and soil great represented by pH in this farmer-friendly CR machine offered study. These fundamental elements are constant indicators across many datasets and significantly influence crop tips. Because of our version's result, the important crop type serves. The suggested observation is typically focused on the quality, vitamins, and health of the soil. The explanations are:

Data Availability & Consistency:

Dataset is a collection of data from multiple sources, most of which are focused on crops, pH, N, P, and K. Certain datasets did include variables like temperature and rainfall, but uniform features across all data points were necessary to preserve consistency and prevent the introduction of data biases.

Complexity in Data Collection for Other Factors:

It is an enormous undertaking to compile a complete dataset that takes into account all climatic and geographic aspects. It involves lengthy data curation procedures, expert interventions, and the possibility of human mistake.

Interrelation with pH:

The pH of the soil, which indicates how acidic or alkaline it is, can serve as a stand-in for a few environmental variables. For example, regular rainfall can affect the pH of the soil, and the pH of the soil can reflect temperature and humidity. The model indirectly accounts for some of the effects of external factors on the soil by taking pH into account.

A Novel Model Training

An AI&ML (AI & ML) model is constructed in Python here. It suggests the optimum crop to plant by analyzing the soil's composition, pH, rainfall, and location. The model focuses on predicting agricultural production by analyzing variables including district, state, season, and crop type using several supervised ML techniques. This method's primary objective is to inform farmers about the crop yield beforehand in order to provide them with crop recommendations that will help them select a crop that will yield more profitably.

The information needed for processing by the ML algorithms is stored in the suggested system. In order to prepare those algorithms for prediction analysis data, this module primarily focuses on testing and training the models using the dataset that was created in the starting module. Following the creation of each analysis, the models are preserved in conversation files for future use in making predictions when their accuracy and validation scores are examined. Additionally, a variety of input types are used to predict various outputs and verify their accuracy.

Data Reduction

ALGORITHM 1 Missing value handling

The performance of the ML models is negatively impacted by garbage, redundant, and unnecessary data (Royston *et al.*, 2006; Benjelloun *et al.*, 2007). We eliminate garbage values, duplicates, and extraneous information to provide the finest ML learnable dataset.

Input: Raw data (S)

Output: Pre-processed dataset after missing value handling

procedure MissingValueHandling(S)

for each attribute Sa do

ma = mean (Sa) [ma is the arithmetic mean of attribute Sa]

for each sample data Sd a do

if Sd a is missing then

Sd a: =ma

```
        end if
    end for
    end for
end procedure
```

A validation set, comprising 10% of the training dataset, is also considered in the study to ensure that the model is not overfitting or underperforming during training. Training data thus consists of one response variable (major crop), four predictors (N, P, K, and pH), and 80% of the data samples. Because soil factors and crop kinds interact in a complex and nonlinear way, the study employed a neural network, a subset of ML models that are adept at seeing intricate patterns and correlations in the data.

AI&ML based Farmer's Friendly Soil and Fertilizer Recommendation System

A specialist tool designed to help farmers make educated judgments about the appropriate amounts of fertilizer to apply to their crops is a farmer-friendly soil and fertilizer recommendation system. The goal of this approach is to minimize negative environmental effects while boosting agricultural yield. A lot of farmers might not be completely aware of the nutrient levels in their soil right now. It is challenging to identify which nutrients are at their optimal levels, which are insufficient, and which are abundant without soil testing. Professional soil testing is expensive, especially for small-scale farmers in poor countries.

The infrastructure required to understand and act upon the results may be lacking, even in situations when testing are fairly priced. A rule-based fertilizer recommendation system is proposed in this study to help farmers choose the right types and amounts of fertilizer for specific crops. The system's foundation is made up of proven scientific ideas from plant biology and soil chemistry. By providing accurate fertilizer recommendations based on these factors, the technique seeks to bridge the knowledge gap because soil pH influences the nutritional needs of different crops.

Numerous elements are considered by the fertilizer recommendation system, such as crop type, the particular nutrients needed for each crop, and the pH level of the soil to determine the quality of the soil. The primary nutrients that crops need are N, P, and K, and the system evaluates these factors to identify the optimal concentrations of each. Because it influences nutrient solubility, which has a significant impact on plant development and optimal production, soil pH is an important characteristic. The pH of the soil can affect the crop's resistance to disease, productivity, and overall health. The pH ranges that are preferred by various crops vary. Using their pH test, the study first creates a lookup table to determine the properties of the soil, as indicated in the table below.

Table: Soil Quality Depending Upon Different Analysis pH Values.

pH Value	Soil Quality
<4.5	Strongly acidic
4.5–5.5	Highly acidic
5.6–6.5	Moderately acidic
6.6–7.0	Slightly acidic
7.0	Neutral
7.1–8.0	Slightly alkaline
8.1–9.0	Moderately alkaline
9.1–10.0	Strongly alkaline
>10.0	Very strongly alkaline

The primary input sources for the rule-based system in the suggested system, which creates a set of logical requirements, are the pH values and associated quality indicator. The project also focuses on the building reference database, which offers recommended ranges of fertilizer (N, P, and K) and suitable pH for different crops. Table 3 displays a sample visualization of the pH and fertilizer recommendations for the crops in question. The information required to create fertilizer recommendations is shown in Table 3. Performing a thorough soil study is the first step. Understanding the unique N-P-K requirements of each crop at various growth stages is crucial since different crop varieties have varied nutritional needs.

The system establishes criteria based on established scientific facts and expert knowledge to ensure reliable recommendations. The system then compares the user's input with fertilizer data to determine whether the pH of the soil is within the permissible range for the chosen crop. The system recommends the optimal levels of N, P, and K based on the fertilizer table's nutrient requirements if the soil's pH is within the acceptable range; if not, it recommends possible soil amendments and prompts the user to alter the soil's pH. Other crops that might thrive in the current soil pH range are also suggested by the algorithm, in addition to recommendations for N, P, and K fertilizer for particular crops. This method eliminates the need for explicit predictive model training on a dataset and enables the system to provide thorough and accurate soil and fertilizer recommendations according to the table below.

Table: pH ranges and Nutrient levels for better crop

Crop	Nitrogen (N)	Phosphorus (P)	Potassium (K)	Suitable pH Range
Maize	120–150	60–80	40–60	5.5–7.0
Sorghum	100–120	50–70	40–60	5.5–7.5
Cassava	60–90	30–40	50–70	5.0–6.5
Beans	60–90	20–30	30–40	5.5–7.0
Potato	80–100	40–60	60–80	5.0–6.5
Coffee	60–80	20–30	40–60	5.5–6.5
Banana	100–120	60–80	60–80	5.0–7.0
Kidney beans	60–90	20–30	30–40	5.5–7.0
Onion	40–60	20–30	30–40	6.0–7.0

Proposed System Model

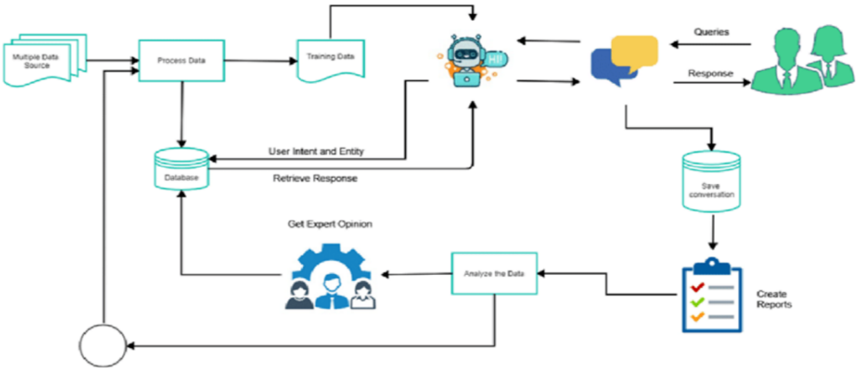

Figure: Farmer's Friendly Soil Fertiliser Recommendation System through Expert Analysis

Crop and Fertilizers Recommendation System through Expert analysis

Figure 1 above illustrates the suggested system design, which takes a highly coordinated approach to computational intelligence and data analytics. Data collection from various sources is the responsibility of the system design's top layer. In the following layer, these data are mapped and examined to create a final crop dataset. After that, correlation analysis and exploratory data analysis are carried out to comprehend the nature of the data and gain important insights regarding the preprocessing methods that should be used to guarantee the dataset's completeness. This process is crucial for making sure the dataset is appropriate for training the learning model for crop recommendation. The study divided the dataset in an 80:20 ratio between training and testing sets, as is normal practice. We set up and improved the neural network model for the specific problem and input data. After that, the trained model was validated using the testing dataset, which contained a number of soil characteristics.

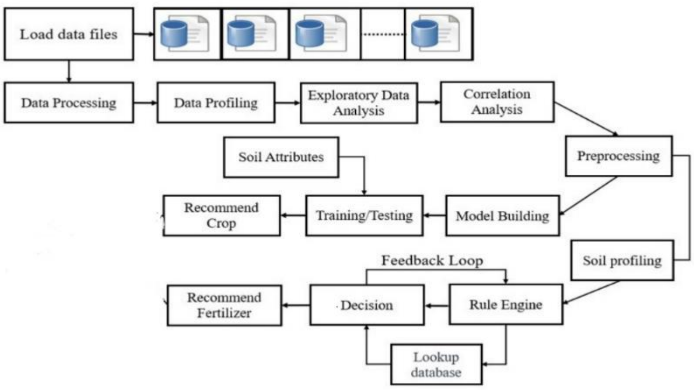

Figure: Proposed Model's Schematic Architecture

The next part of the proposed system incorporates a rule-based fertilizer recommendation system by first profiling the soil using the pre-processed dataset from the crop recommendation system and then generating a lookup table based on scientific facts and expert knowledge. The underlying philosophy of the proposed approach is that, while ML can be helpful for many tasks, domain-specific, expert-driven rules are undeniable.

The proposed rule-based fertilizer recommendation method is based on the idea that each crop and soil combination have distinct nutritional requirements, and it is easy to understand because it is based on established agricultural data. This enables farmers to comprehend the rationale behind the suggestions, fostering trust and potentially resulting in increased agricultural yields and more environmentally friendly farming methods. A feedback loop mechanism is also incorporated into the system to assist in updating and improving the guidelines and suggestions over time. In addition to discussing the system's basic architecture, the proposed study emphasizes the system's practicality and viability for cloud implementation.

Table: Features Mean Values for Each Crop

Index	Crop Name	Nitrogen	Phosphorous	Temperature	Humidity	pH	Rainfall
1	rice	79.89	47.58	23.69	82.27	6.43	236.18
2	maize	77.76	48.44	22.39	65.09	6.25	84.77
3	chickpea	40.09	67.79	18.87	16.86	7.34	80.06
4	kidneybeans	20.75	67.54	20.12	21.61	5.75	105.92
5	pigeonpeas	20.73	67.73	27.74	48.06	5.79	149.46
6	mothbeans	21.44	48.01	28.19	53.16	6.83	51.20
7	mungbean	20.99	47.28	28.53	85.50	6.72	48.40
8	blackgram	40.02	67.47	29.97	65.12	7.13	67.88
9	lentil	18.77	68.36	24.51	64.80	6.93	45.68
10	pomegranate	18.87	18.75	21.84	90.13	6.43	107.53
11	banana	100.23	82.01	27.38	80.36	5.98	104.63
12	mango	20.07	27.18	31.21	50.16	5.77	94.70
13	grapes	23.18	132.53	23.85	81.88	6.03	69.61
14	watermelon	99.42	17.00	25.59	85.16	6.50	50.79
15	muskmelon	100.32	17.72	28.66	92.34	6.36	24.69
16	apple	20.80	134.22	22.63	92.33	5.93	112.65
17	orange	19.58	16.55	22.77	92.17	7.02	110.47
18	papaya	49.88	59.05	33.72	92.40	6.74	142.63
19	coconut	21.98	16.93	27.41	94.84	5.98	175.69
20	cotton	117.77	46.24	23.99	79.84	6.91	80.40
21	jute	78.40	46.86	24.96	79.64	6.73	174.79
22	coffee	101.20	28.74	25.54	58.87	6.79	158.07

Performance Metrics and Discussion

The recommended recommendation models are built using the Python programming language via the Anaconda software, which is installed on a Windows 10 PC. In this part, the performance analysis of the proposed crop recommendation prediction model is displayed. Metrics such as the ROC curve, F1-score, recall, accuracy, precision, and recall were employed to assess the accuracy of the predictions. These metrics ensure the accuracy and reliability of the model's recommendations by providing a comprehensive evaluation of its performance. The performance metric accuracy is the proportion of all predictions that were correct. This is how it is supplied:

$$\text{Accuracy} = TP + TN / TP + TN + FP + FN$$

In which, FP represents False Positives, FN represents False Negatives, TP represents True Positives, and TN represents True Negatives. The percentage of positive identifications that were truly accurate is known as precision. This is how it is calculated:

$$\text{Precision} = TP / TP + FP$$

The capacity of a model to identify every pertinent instance in a dataset is known as recall (or sensitivity). The recall is provided by:

$$\text{Recall} = TP/TP + FN$$

The F1 Score seeks to strike a balance between precision and recall by taking the harmonic mean of the two. The following formula is used to calculate the F1 Score:

$$\text{F1 Score} = 2 \times \text{Precision} \times \text{Recall} / \text{Precision} + \text{Recall}$$

Table 1: Chemical properties of fertilized plot with different types of organic manure

Parameter	Initial soil analysis[a]	Treatments[a]			
		p1	p2	p3	p4
Organic carbon (%)	0.82	2.07	2.37	1.98	2.34
Nitrogen (%)	0.10	0.20	0.25	0.23	0.23
C/N ratio	8.00	11.00	10.00	9.00	10.00

[a] Mean of two replicates.

Performance Analysis

The above table shows the average NPK (kg/ha) and pH requirements for different crops. On average, bananas need 100.19 kg of nitrogen, 80.89 kg of phosphorus, and 50.04 kg of potassium per hectare. Banana soil has an average pH of 6.07, which is regarded as slightly acidic. On average, beans require 75.09 kg of nitrogen, 25.32 kg of phosphorus, and 34.73 kg of potassium per hectare. Beans require an average pH of 5.89, which is regarded as being very acidic. A hectare of cassava typically requires 59.79 kg of potassium, 34.92 kg of phosphorus, and 74.77 kg of nitrogen. The average pH of cassava fields is 5.92, which is rather acidic. Similar interpretations can be created for many crops to understand their fertilizer requirements and soil compatibility for more efficient crop production.

Table: Data visualization of crops soil data analysis

	Major_CROP	Nitrogen	Phosphorus	Potassium	pH
0	banana	110	70	70	6.00
1	beans	80	25	35	6.25
2	cassava	135	70	50	6.24
3	coffee	70	25	50	6.00
4	kidneybeans	75	25	35	6.26
5	maize	135	70	50	6.26
6	onion	50	25	35	6.49
7	potato	90	50	70	5.75
8	sorghum	110	60	50	6.48

The following figure compares several crops according to their Kg/ha nitrogen requirements. According to the data, sorghum, onions, and maize are the crops that need nitrogen the most. The least nitrogen-intensive plant is kidney beans. Bananas can be classified as having both the highest and lowest nitrogen demands. Comparing various crops according to their phosphorus requirements in kilograms per hectare is also shown in the Figure 3.

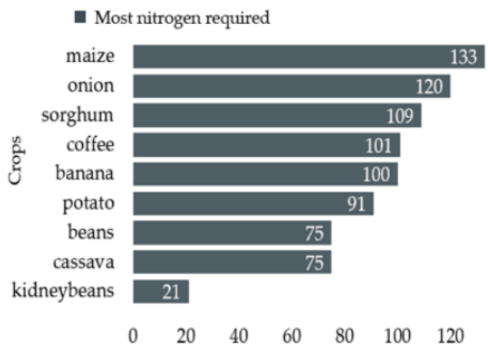

Figure: Analysis regarding most and least nitrogen-requiring crops.

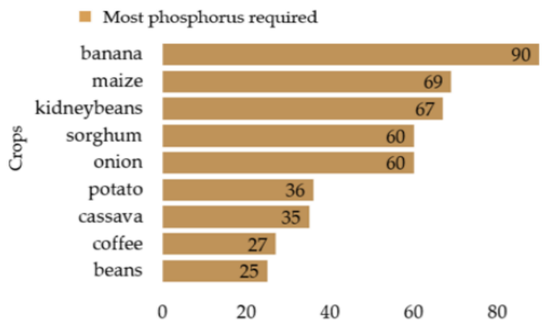

Figure: Analysis concerning most and least phosphorus-requiring crops.

The aforementioned chart indicates that bananas require the highest amount of phosphorus, followed by kidney beans and maize. Beans and coffee require less phosphorus than other foods. Onions are among the top categories and require different amounts of phosphorus. Onions, cassava, and potatoes have the highest potassium requirements, whereas kidney beans, maize, and sorghum have the lowest, according to a closer look at the chart below. The differing nutritional profiles of several crops across categories are highlighted in this analysis.

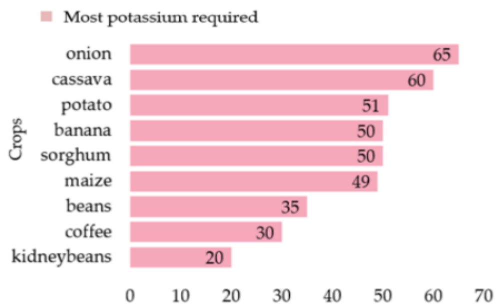

Figure: Analysis for most and least potassium-requiring crops.

A ML model's training phase is essential since it establishes how well the model recognizes patterns in the training data. The model's ability to predict the labels of the training data is gauged by its training accuracy. A model with low training accuracy may have issues with the hyperparameters, the model architecture, or the dataset. Another crucial parameter for making sure the model generalizes well that is, that it can produce precise predictions on previously unseen data is training accuracy. This is due to the fact that a model that achieves high training accuracy after achieving high validation accuracy has discovered the basic patterns in the data rather than simply the particular instances in the training set.

Scope and Limitations

The proposed rule-based fertilizer system and crop recommendation has a lot of potential because of its wide range of applications. Using soil data properties and ML techniques, it considers N, P, K, and pH of the soil. Depending on the state of the soil, these elements are critical to crop development. A controlled pH ensures proper nutrient uptake, which supports healthy crop growth. Ensuring that these crucial characteristics are within their ideal range's accounts for a significant portion of crop health following fertilizer suggestion yield improvement. The method guarantees a customized approach to crop cultivation, based on the unique circumstances and requirements of the soil, by taking into account certain crucial soil characteristics. By concentrating on these four crucial factors, the suggested approach has drawbacks when it comes to widespread adoption in the agricultural industry, even with its benefits. The following highlights the possible difficulties and restrictions.

Data Generalization: Developing a complete dataset that encompasses all potential soil, agricultural, environmental, and geographic factors is challenging. Inconsistent or insufficient data can also result in errors.

Infrastructure Challenges: Extensive infrastructure, such as data transmission networks and Internet of Things sensors, may be needed for large-scale deployment.

Maintenance and Updating: As crop types, agricultural methods, and soil conditions vary, the model will need to be updated on a regular basis. This will be difficult on a big scale.

Conclusion

This study's new Farmer's Friendly Soil Fertilizer recommendation system was developed through expert evaluation. Using artificial intelligence and ML, the system uses soil data as its primary basis to give farmers information that could help them choose crops and apply fertilizer more intelligently. In extensive comparisons, the neural network outperformed a number of popular ML models, demonstrating its precision, reliability, and skill. Along with reducing the environmental impact and promoting economical agricultural practices, the system might also boost crop yield and penalties. But the device has several limitations. The process of data modelling and function extraction needs to be delicate and multidimensional, and more regional and environmental elements should be taken into account. A suitable crop is selected, and the outcomes are carefully evaluated. The observer determines that XGBoost and Naive Bayes are the most suitable, accurate, and eco-friendly classifiers. universal, this research has been extremely beneficial to the agricultural field. The method's accuracy, scalability, and ease of use make it a useful tool for farmers, governments, and organizations.

Future work

It is possible to expand this work in several ways. Some instances are shown below. Any of the following ideas are available for readers to build upon in this work. 1) Survey farmers to learn how much they save with the technique. That way, the economic impact of certain ML models may be ascertained. 2) A mobile application that serves as an end-to-end tool for potential customers (farmers or agribusiness owners) could be created in addition to this endeavor. Specifically, using the models covered in the paper to expand a business proposition. This will make it more likely that farmers and other users will be able to use these methods. 3) collect data from specific areas, which will enable us to determine whether the techniques are reliable in exceptional situations. 4) Utilize a bigger dataset; the more data we have, the better the models can examine the relationships between various factors and agricultural yields. 5) Analyse the financial and environmental impacts to determine whether the strategy has any potential advantages for both farmers and the environment. 6) Sensors can be installed on the farm to gather information that can be used to enhance agricultural advisory systems, provide real-time forecasts, and reduce crop loss risk. Farmers will gain from more sustainable growth, greater selection (due to real-time data), and reduced labor expenses.

References

[1]. IBM Research India. (2020). AgroAI: AI-Powered Agriculture. https://www.research.ibm.com/artificialintelligence/projects/agro-ai.

[2]. Parvathala, Balakesava & Manikandan, A. & Vijayalakshmi, P. & Parvez, M. & Gopalan, S. & Ramalingam, S.. (2024). Bio-Inspired Metaheuristic Algorithm for Network Intrusion Detection System of Architecture. 10.4018/979-8-3693-5276-2.ch004.

[3]. Zhang, Y., Wang, L., Huang, W., & Guo, Y. (2020). An intelligent system for crop management based on AI and IoT technologies. Sensors, 20(18), 5245.

[4]. Ali, R., Manikandan, A., Lei, R. et al. A novel SpaSA based hyper-parameter optimized FCEDN with adaptive CNN classification for skin cancer detection. Sci Rep 14, 9336 (2024). https://doi.org/10.1038/s41598-024-57393-4.

[5]. Harihara Gopalan, S., Muzammil Parvez, M., Manikandan, A., & Ramalingam, S. (2024). Cognitive radio spectrum allocation using Nash equilibrium with multiple scheduling resource selection algorithm. Ain Shams Engineering Journal. https://doi.org/10.1016/j.asej.2024.102688.

[6]. Pawlak, K.; Kołodziejczak, M. The Role of Agriculture in Ensuring Food Security in Developing Countries: Considerations in the Context of the Problem of Sustainable Food Production. Sustainability 2020, 12, 5488.

[7]. Kolli, Srinivas & V., Praveen & John, Ashok & Manikandan, A.. (2023). Internet of Things for Pervasive and Personalized Healthcare: Architecture, Technologies, Components, Applications, and Prototype Development. 10.4018/978-1-6684-8913-0.ch008.

[8]. Palaniappan, Mathiyalagan & Annamalai, Manikandan. (2019). Advances in Signal and Image Processing in Biomedical Applications. 10.5772/intechopen.88759.

[9]. Ali, R., Manikandan, A. & Xu, J. A Novel framework of Adaptive fuzzy-GLCM Segmentation and Fuzzy with Capsules Network (F-CapsNet) Classification. Neural Comput & Applic (2023). https://doi.org/10.1007/s00521-023-08666-y.

[10]. V. A.R, S. David, E. Govinda, K. Ganapriya, R. Dhanapal and A. Manikandan, "An Automatic Brain Tumors Detection and Classification Using Deep Convolutional Neural Network with VGG-19," *2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, Coimbatore, India, 2023, pp. 1-5, doi: 10.1109/ICAECA56562.2023.10200949.

[11]. Cisternas, I.; Velásquez, I.; Caro, A.; Rodríguez, A. Systematic literature review of implementations of precision agriculture. Comput. Electron. Agric. 2020, 176, 105626.

[12]. Annamalai, Manikandan & Muthiah, Ponni. (2022). An Early Prediction of Tumor in Heart by Cardiac Masses Classification in Echocardiogram Images Using Robust Back Propagation Neural Network Classifier. Brazilian Archives of Biology and Technology. 65. 10.1590/1678-4324-2022210316.

[13]. Hossain, M.A.; Siddique, M.N.A. Online Fertilizer Recommendation System (OFRS): A Step Towards Precision Agriculture And Optimized Fertilizer Usage By Smallholder Farmers In Bangladesh: Online fertilizer recommendation. Eur. J. Environ. Earth Sci. 2020, 1.

[14]. Manikandan, Annamalai, M,Ponni Bala. (2023). Intracardiac Mass Detection and Classification Using Double Convolutional Neural Network Classifier. Journal of Engineering Research. 11(2A). 272-280. 10. 36909/jer.12237.

[15]. Singh, R.K.; Berkvens, R.; Weyn, M. AgriFusion: An Architecture for IoT and Emerging Technologies Based on a Precision Agriculture Survey. IEEE Access 2021, 9, 136253–136283.

[16]. Balamurugan, D. & Seshadri, s.Aravinth & Reddy, P. & Rupani, Ajay & Manikandan, A.. (2022). Multiview Objects Recognition Using Deep Learning-Based Wrap-CNN with Voting Scheme. Neural Processing Letters. 54. 1-27. 10.1007/s11063-021-10679-4.

[17]. Available online: https://www.nri.org/images/documents/development-programmes/gender_soc_dif/publications/farming_ for_impact_-_full_report.pdf (accessed on 18 September 2023).
[18]. Rivera, J.I.; Bonilla, C.A. Predicting soil aggregate stability using readily available soil properties and ML techniques. Catena 2020, 187, 104408.
[19]. Suchithra, M.S.; Pai, M.L. Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters. Inf. Process. Agric. 2020, 7, 72–82.
[20]. Chambers, O. ML Strategy for Soil Nutrients Prediction Using Spectroscopic Method. Sensors 2021, 21, 4208.
[21]. Rose, S.; Nickolas, S.; Sangeetha, S. ML and Statistical Approaches used in Estimating Parameters that Affect the Soil Fertility Status: A Survey. In Proceedings of the 2018 Second International Conference on Green Computing and Internet of Things (ICGCIoT), Karnataka, India, 16–18 August 2018; IEEE: New York, NY, USA, 2018; pp. 381–385.
[22]. Rajamanickam, J. Predictive model construction for prediction of soil fertility using decision tree ML algorithm. Infocomp J. Comput. Sci. 2021, 20, 49–55.
[23]. Klerkx, L.; Jakku, E.; Labarthe, P. A review of social science on digital agriculture, smart farming and agriculture 4.0. New contributions and a future research agenda. Njas-Wagening. J. Life Sci. 2019, 90, 100315.
[24]. Talaviya, T.; Shah, D.; Patel, N.; Yagnik, H.; Shah, M. Implementation of artificial intelligence in agriculture for optimization of irrigation and application of pesticides and herbicides. Artif. Intell. Agric. 2020, 4, 58–73.
[25]. Priya, P.K.; Yuvaraj, N. An IoT Based Gradient Descent Approach for Precision Crop Suggestion using MLP. J. Phys. Conf. Ser. 2019, 1362, 012038.
[26]. Gupta, A.; Nagda, D.; Nikhare, P.; Sandbhor, A. Smart Crop Prediction using IoT and ML. Int. J. Eng. Res. Technol. 2021, 9, 18–21.
[27]. Vivekanandhan, V.R.; Sakthivel, S.; Manikandan, M. Adaptive neuro fuzzy inference system to enhance the classification performance in smart irrigation system. Comput. Intell. 2021, 38, 308–322.
[28]. Sharma, S. Crop Recommendation Dataset [Data Set]. 2021. Available online: https://www.kaggle.com/datasets/siddharthss/crop-recommendation-dataset (accessed on 18 September 2023).
[29]. Jaiswal, R. Soil Fertility Dataset [DataSet]. 2023. Available online: https://www.kaggle.com/datasets/rahuljaiswalonkaggle/soil-fertility-dataset (accessed on 18 September 2023).
[30]. Perez-Guzman, K.; Imanirareba, D.; Jones, S.K.; Neubauer, R.; Niyitanga, F.; Naramabuye, F.X. Sustainability implications of Rwanda’s Vision 2050 long-term development strategy. Sustain. Sci. 2023, 18, 485–499.