PRECISION FARMING UTILIZING MACHINE LEARNING FOR CROP AND FERTILIZER RECOMMENDATIONS

Allwyn Rajkumar M[1], Mr.V.Balamurugan[2], Mrs.M.S.Kavitha[3], Mrs.M.Nisha[4]
[1]PG Scholar,[2,3,4] Assistant Professor
Department of Computer Science and Engineering
Akshaya College of Engineering and Technology, Bhagavathipalayam, Kinathukadavu, Coimbatore, Tamil Nadu 642109

Abstract

Predicting crop yields is a challenging and time-consuming process in agriculture. Decisions on a global, regional, and field scale rely on it heavily. Predicting agricultural productivity is one example of a mosaic characteristic that is influenced by several variables, including genetics, environment, and interactions between the two. A basic familiarity with the components that affect yield is necessary for accurate yield prediction. Strong algorithms and large datasets are required to discover the connection between the linked components. Machine learning is an essential technique for agricultural yield forecasting since it allows for educated decisions on crop planting and growing season care. Farmers are better able to plan investments and anticipate returns with the use of an accurate crop prediction model. The "Crop Prediction with Machine Learning" project was built using the front-end languages of HTML, CSS, and JS and the back-end language of PYTHON. Agriculture provides a living for the vast majority of people in countries like India. The agricultural sector is seeing the widespread use of cutting-edge technology like Machine Learning while Deep Learning to facilitate increased crop yields with less effort on the part of farmers. To aid farmers in making more informed choices, the "Crop Prediction & Fertilizer Recommendation Applying Machine Learning" initiative has introduced a website. The website would not function without the Crop Recommendation while Fertilizer Recommendation functionalities. Farmers may benefit from Crop Recommendation by learning which crops thrive in their specific environments. This aids them in cultivating appropriate crops and making efficient use of their land. Farmers may find out what nutrients their soil requires to grow their crops in Fertilizer Recommendation. This will allow them to make better use of the fertilizers and cut down on waste. Even farmers without much experience with technology will find the website straightforward to use. Improved sustainability, resource management, and environmental protection are all outcomes of the usage of cutting-edge AI technology, which aids farmers. The long-term goal of this effort is to improve farming.

Keywords— Crop Prediction; Feature Selection; Random Forest; Machine Learning; Smart Farming

I. INTRODUCTION

The introduction of precision agriculture has changed the game for farmers, and new models for recommending crops and fertilizer have emerged as a result. Soil nutrient phases, moisture content, and pictures are just a few of the data sources used by these models to optimize fertilizer usage and crop yields. A lack of data, high-priced sensors and other technology, and a requirement for specialized knowledge to develop and oversee different solutions are some of the obstacles to machine learning's use in agriculture. Nevertheless, the prospective benefits of using machine learning in farming will become more apparent as an increasing number of farms adopt precision agriculture and collect data. Machine learning as it pertains to agriculture is still in its early stages, therefore more research is needed to fully use its potential. The use of machine learning is only likely to increase in the coming years, and early findings are promising. Precision farming stands out as an example of machine learning in agriculture. This method optimizes agricultural practices including fertilization, irrigation, with pest control using data and technology, leading to higher yields and better quality of harvest. Aerial footage, soil sensors, and satellite photos are just a few of the numerous sources of data that machine learning algorithms can sort through to provide precise maps of crop growth, nutrient concentrations, as well as moisture levels. In order to optimize harvest yield and avoid waste, farmers may use these maps to control agricultural behaviors like watering certain field sections or adding fertilizer. Using machine learning methods, massive amounts of data may be sorted through, including data collected from IoT devices and other sources. The area is expanding at a fast pace and may revolutionize agricultural production forecasting and analysis. Machine learning algorithms allow computers to learn and become better over time without human intervention by analyzing data using mathematical and statistical models and algorithms. Information specific to agricultural production, such as weather patterns, soil properties, crop growth stages, and outbreaks of disease and pests, may teach machine learning algorithms a great deal. By analyzing the collected data, machine learning algorithms may make very precise forecasts about growth, quality, and production. Most current crop recommendation algorithms only take into account one kind of data, such soil nutrient levels or weather, when making their predictions. Despite these models' promise for improving crop yields, their reliance on Interpretable Long Short-Term Memory Networks (ILSTM) causes them to overlook important details about the agricultural environment, such as the interplay between various data kinds [3, 4], and [5]. Lacking appropriate data integration techniques, several models have also tried to use multimodal information to crop suggestions, but their capabilities are restricted [6], [7]. Current models for fertilizer recommendations, similar to those for crop recommendations, often depend on just one kind of data. Soil nutrient levels have been used to train models that, using the Phenology Normalization and Deep Learning (PNDL) method, can advise farmers on what kind and how much fertilizer to apply to various crops [8, 9]. In terms of maximizing the use of fertilizer, such models have shown promise. But since they don't take into consideration the synergistic impacts of several data kinds, they often fail to provide correct suggestions [10], [11], [12]. Our suggested model will have a major effect on precision agriculture.

Revolutionizing precision agriculture, the integration of multimodal data with modern data processing methods may provide farmers with personalized advice to maximize crop yields with fertilizer usage. As a result, farmers may maximize yields with minimal inputs like fertilizers, which improves sustainability and reduces environmental impact. Finally, our suggested paradigm is a huge step forward for precision farming. Improving results by including multimodal data, creative data processing methods, and unique models into our approach shows its potential as an essential instrument for precision agriculture. Our efforts to advance precision agriculture are significant because of the potential benefits to sustainability while the environment, as well as the enhancements to crop while fertilizer recommendation systems. The suggested model is contrasted with the most recent published research based on a number of parameters in table 1.

II. RELATED WORK

Using network operations such as Wavelet-attention convolutional neural networks (WACNNs) and others to provide accurate and thorough recommendations is currently the biggest challenge for existing models [13], [14], [15]. It often leads to less-than-ideal suggestions that don't always take into account the specifics of the agricultural environment. While developing personalized recommendations with 3D-Convolutional Neural Networks and Attention Convolutional LSTM Methods (3DCNN ACLSTM) [16], it is essential to incorporate temporal dependencies and complex interactions between different data sources. These aspects are often ignored by existing models. The possibility of using multimodal data to inform fertilizer and crop recommendations has been investigated by a number of researchers. The context-aware fertilizer recommendation (CAFR) procedure is one example of an effective data fusion tool that has been lacking in these endeavors [17], [18], [19]. Create precise and all-encompassing ensemble LSTM (eLSTM) model recommendations by integrating varied data types such NPK levels, moisture content, imaging, and geographical information [20], [21]. For suggestions to be informed by significant patterns and correlations, multimodal data analysis is crucial. The complexity of multimodal sample data may be too much for existing models to handle, as they depend on conventional data processing techniques [22]. To the contrary, our suggested model makes use of cutting-edge data processing methods to glean useful information from the data samples [23], [24], including frequency pattern analysis, entropy design analysis, S Transform components extraction, and convolutional elements analysis. A major step forward in precision agriculture is our suggested concept. Through the integration of multimodal data sources, our model offers a comprehensive perspective of the agricultural environment. The novel data processing methods used by our model enable the extraction of significant connections and patterns [25], [26]. Recommendations are even more accurate and personalized because to the BiGRU features, ALFPCA for feature selection, along with the GCFPMax and RFPMax models [27]. By precisely capturing time-dependent and complex interactions between various data kinds, our model also overcomes the shortcomings of previous approaches [28].In terms of accuracy, precision, and memory, comparative investigations demonstrate that our suggested model surpasses current models [3].

We found that our model significantly improved crop and fertilizer recommendation accuracy by 4.9%, recall by 2.5%, and precision of crop recommendations with 3.5%.

III. PROBLEM STATEMENT

Under the current setup, farmers mostly use their gut feelings and long-held knowledge of the area when deciding what crops to grow, how much fertilizer to use, and how to deal with diseases. Farmers may have difficulties in determining which crops are most suited for this method, and they may also have trouble effectively regulating their nutrients. Agricultural sustainability and production might be hindered by a lack of accurate information and innovative technology.

Problems

- Conventional approaches devoid of insights derived from data.
- Wasted resources because cutting-edge technology is nonexistent.
- Farmers cling to what they know, which means fewer crop types.
- Delays in illness detection, leading to the spread of the disease.
- Losses in crops pose a greater threat to farmers' incomes and the availability of food.
- The effects on the environment of inefficient chemical and nutrient management.

IV. PROPOSED WORK

Minimizing water use for agricultural production and increasing crop yield are both achieved via the effective operational oversight of drip irrigation. The use of drip irrigation on farmlands for farming, particularly for high-cost commodities like fruits and vegetables, has been much praised for its social, economic, and environmental benefits. In addition, the low-pressure watering mechanism that drip irrigation is built on makes it more energy efficient than sprinkler systems. The "Crop prediction using Fertilizer Recommendation using Machine Learning" proposal is a groundbreaking effort to transform farming by using cutting-edge machine learning technologies. Crop recommendation, fertilizer recommendation, and this online platform are the three components it offers. Improving agricultural production and encouraging sustainable practices that minimize ecological damage and resource depletion are the goals of the initiative, which seeks to equip farmers with data-driven insights. Optimizing resource usage and enhancing output potential, the Crop Recommendation feature helps farmers choose the best crops for their soil conditions. Also, the Fertilizer Recommendation component analyzes soil data and gives you choices for fertilizer based on your specific needs, so you don't waste any and the environment stays protected.

DATASET PREPARATION

Acquiring data from IoT devices on farms is crucial for machine learning analysis. Collecting crucial data on agricultural activities, such as crop type, water demands, harvesting procedures, and crop usage, may significantly improve performance on smart farms. The dataset collection for this component should contain nutrients that include nitrogen, phosphorus, potassium, temperature, humidity, and rainfall, as well as pH, temperature and humidity. You may thank Kaggle for providing these datasets. Rice, maize, beans, moth beans, mung beans, lentils, pigeon peas, watermelon, muskmelon, orange, papaya, coconut, cotton, pomegranate, banana, mango, grapes, jute, coffee—that's the 22-item count. The 22 Goal classes that our model was trained on are these.

A grand total of eight traits are considered during the training process. Some examples of these features include precipitation, humidity, temperature, pH, nitrogen, phosphorus, and potassium. The Kaggle database, an online hub for researchers to exchange and access data from various projects, is used to compile the data. Among the many parameters included in the dataset are the following: soil temperature, pH, rainfall, humidity, phosphorus to potassium, and nitrogen to temperature ratios. There are 2200 records in the crop forecast dataset with 22 crop labels. Some of these crops include oranges, chickpeas, jute, grapes, watermelon, black gram, watermelon, bananas, rice, cotton, chickpeas, coconuts, and kidney beans. The dataset contains a number of characteristics, including the following: temperature, soil pH, rainfall, humidity, phosphorus content ratio, potassium content ratio, nitrogen content ratio, and other similar measurements. 67 percent of the information is used for training purposes, while the remaining data is utilized for testing.

DATA PREPROCESSING

Experiments including label changes were used to assess the precision of data processing. Instead of trying to forecast specific kinds of crops, we divided them into four categories according to a number of criteria. With the use of machine learning, we classified a dataset of crop varieties based on seven distinct characteristics. On top of that, we calculated how few characteristics are required for accurate learning and prediction. Take care of any problems with the dataset, such as inconsistent data, outliers, or missing values. Consider the data's characteristics when deciding whether to impute or eliminate missing values; deal with outliers in a way that keeps them from impacting the model's performance. Dataset consistency and reliability are guaranteed during cleaning, which is essential for model training.

FEATURE EXTRACTION

Improve the model's prediction capabilities by adding or modifying characteristics.

SI No	Description	Count
1	Total Instances	2200
2	Total Features	8
3	Target Classes	22

Crop Recommendation Summary (Table 1)

When it comes to classification and regression issues, supervised machine learning algorithms like Random Forest come in handy. In order to provide a forecast, this ensemble learning system consults Decision Trees. Random Forest constructs a large number of Decision Trees and takes an average of their forecasts to get a final forecast. Raising the tree density improves accuracy and robustness. This is the pseudocode for the random forest algorithm:

This data set has the following variables: {{number of attributes to evaluate at each split: f}}, a maximum depth of d trees, a training data set with the coordinates {{x, y}}

Outputs: {Learned trees for classification}

- Select a bootstrap sample for each Random Forest tree from the training data set.
- The maximum depth of a Decision Tree T_t should be d.
- At each T_t split, choose f features at random to evaluate.
- For each node of T_t, determine the optimal split using the attributes that were chosen.

Make a table with all the Decision Trees: T_1, T_2,..., T_T.

For every piece of input information:

- Figure out the forecast for every choice.
- Collect forecasts from each tree.
- Determine the final forecasted group.

Model Training

By carefully curating the characteristics that our algorithms for machine learning use, we aim to accomplish accurate crop recognition. Parameter parameters on the necessary qualities were taken into consideration to guarantee optimum outcomes. We were able to get trustworthy information for our farming activities because of this. We tested how different label changes affected the precision of our Analysis of data method. Because of this, we were able to refine our technique and get more accurate findings by better understanding the effect of small label modifications. Use of more generalized labels is critical for accurate crop categorization. Extensive research into this area is required to identify the best categorization methods.

TRAIN-TEST SPLIT

Separate the dataset into a training set and a testing set. Machine learning models are trained using training sets and evaluated using testing sets. A typical distribution is as follows: 80% for training and 20% for testing. The machine learning system may now be trained using the provided dataset. Subsequently, the model may be assessed, fine-tuned, and used for crop suggestion using input characteristics. To enhance the model's accuracy while generalizability, we repeat these stages as necessary, tweaking the dataset preparation procedure.

CROP RECOMMENDATION

Input characteristics including soil type, climate, and previous performance should be used by the model to anticipate appropriate crops.

USER INTERACTION

Make it possible for farmers and agronomists to enter pertinent data using an API or user interface. Make the suggested crops and fertilizers visible, and provide explanations, so everyone can see.

MODEL EVALUATION

If you want to know how well the model worked, you may look at its F1 score, accuracy, precision, and recall. By running the model on previously unknown samples, you may ensure its applicability to fresh data.

V. RESULTS & DISCUSSION

Python is an object-oriented, high-level language with dynamic semantics. The language may also serve as an interpreter. The first step was to gather the data, which was then standardized and missing values were substituted with means. This language is great for quickly creating applications, scripting them, or gluing together pre-existing components for crop recommendation because of its high-level built-in databases, dynamic typing, and dynamic binding. The next step in building the model is to use a random forest classifier. Afterwards, the data is divided into two groups based on the number of classes: regression and classification. An increase in the total amount of trees improves the accuracy of the prediction. We employ a number of different dataset chunks during training. When training the model, only 80% of the information is actually used. The remaining information serves as the foundation for the experiment.

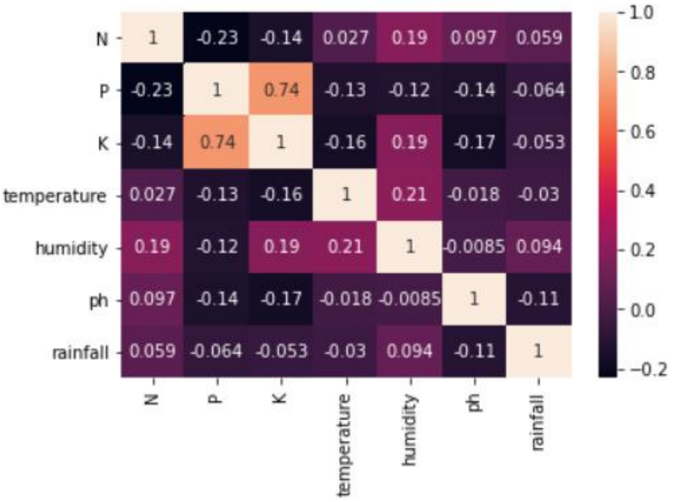
The list of 22 crops includes rice, corn, beans, pigeon peas, grapes, watermelon, muskmelon, apple, moth beans, mung beans, lentils, orange, papaya, coconut, cotton, jute, pomegranate, banana, mango, coffee. The 22 Target classes that our model was trained on are these. During training, a total of eight characteristics are taken into account. Features such as these are labeled, temperature, humidity, pH, nitrogen, phosphorus, and potassium.

HARDWARE CONFIGURATION:

Processor	: Intel icore 7 5th gen
Hard disk	: 500 GB
Ram	: 12 GB
Keyboard	: Logitech of 104 keys
Mouse	: Logitech mouse
Monitor	: 14 inch samtron monitor
GPU	: NVIDIA Geforce GTX 1650

SOFTWARE CONFIGURATION:

Front end Language	: HTML, CSS, Bootstrap, JavaScript : python
Operating system	: Windows 10
Tools	: python IDLE



Heatmap of features

SI No	Description	Count
1	Total Instances	2200
2	Total Features	8
3	Target Classes	22

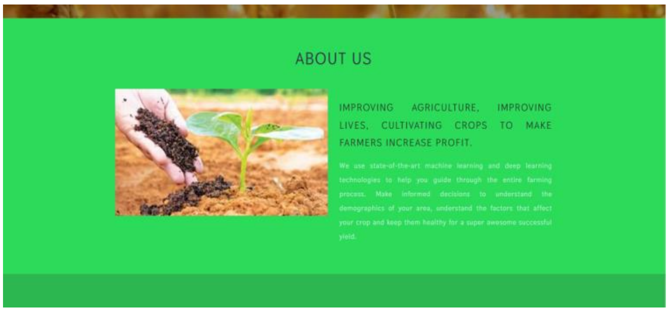
Table 1. Description of crop_recoomedation.csv

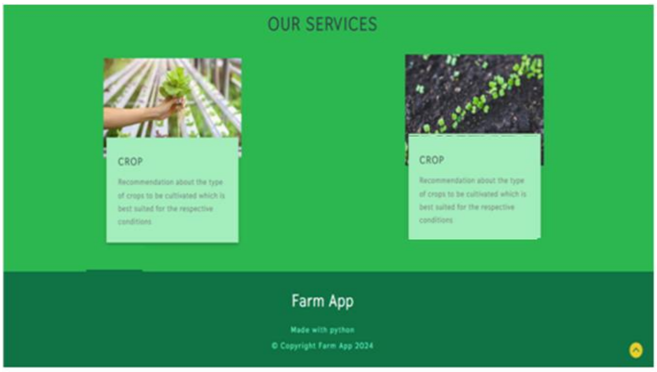
Form name: Home page

Description: The website's user interface is intuitive.

Form name: About Us Page

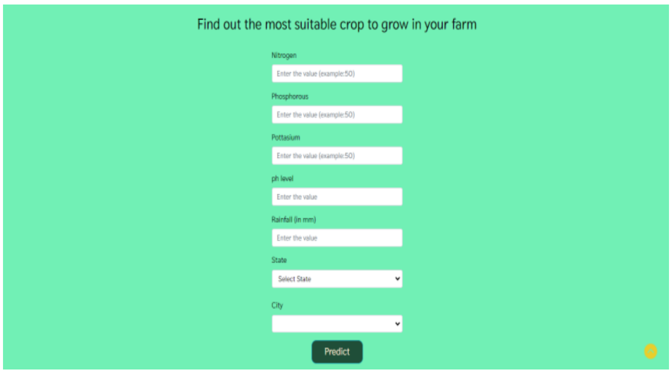
Description: The user view website services.

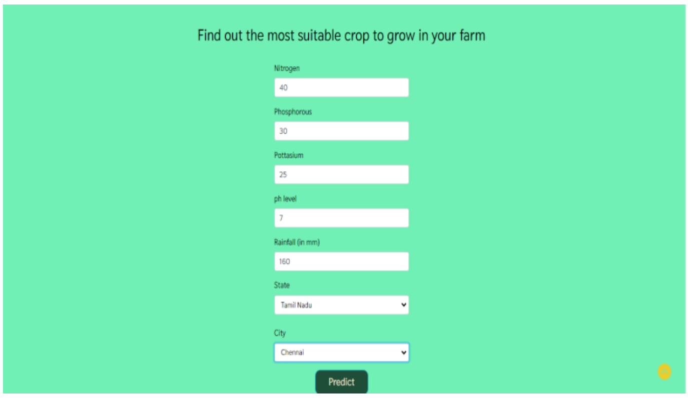




Form name: Crop Recommendation details

Description: The website allows the user to input facts about the soil and location.





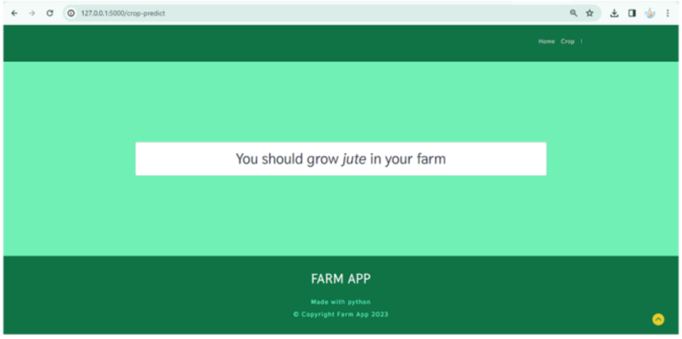
OUTPUT DESIGN

A computerized system's output design is a crucial idea since, without it, the user can think the system is pointless and decide not to use it. Making good decisions is made easier with the right output design, which is critical for every system. Several reports are part of this system's output design.

The most crucial and immediate source of information for the user is the computer's output. The system's relations with the user along with decision-making should be enhanced by efficient and understandable output design. Hard copies produced by printers are a common kind of output.

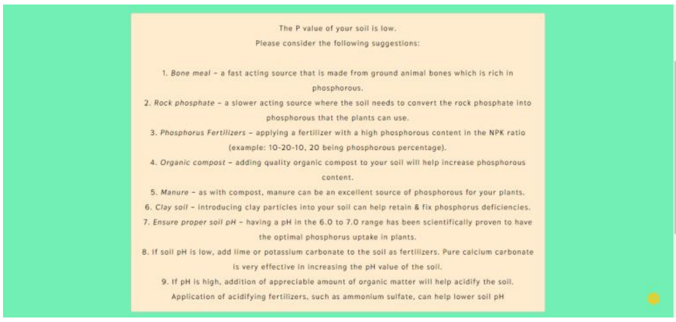
Form name: Crop Recommendation details

Description: the report contains recommended crop details



Form name: Crop Recommendation details

Description: this report includes a fertilizer that is tailored to the specific needs of each plant



The accuracy values while error rates of the algorithms that were studied are shown in Table 1. The algorithms that provide the highest level of accuracy include Bayes Net, Naïve Bayes Classifier, Hoffding Tree, and Random Forest. While random forest classification has an accuracy of above 90%, the DT algorithm achieves 88.50%. Here is a way to formalize error metrics:

$$K(\text{ Kappa }) = (P_0 - P_\mathrm{e})/(1 - P_\mathrm{e})$$

where P_0 as well as P_e, the hypothetical probability of random agreement, denote relative observed agreement.

$$\mid \text{ MAE (Mean Absolute Error) valueI } = \frac{\sum_{\mathrm{I}=1}^{\mathrm{N}} |\mathrm{y_i} - \mathrm{x_i}|}{\mathrm{n}}$$

where n is the whole data set, x_i is the actual value, while $\mathrm{y_i}$is the forecast.

$$\text{RMSE (Root Mean Square Error) } = \sqrt{\frac{[\sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{x}_1 - \mathrm{x}_1')]}{\mathrm{n}}}$$

where x_i is experiential and x_i is the analytical value.

$$\text{RAE (Relative Absolute Error) } = \frac{\sum_{i=1}^{n} |y_i - y_i|}{\sum_{i=1}^{n} |y_i - y'|}$$

where y is the data's mean.

$$\text{Root Relative Squared Error (RRSE) } = A = \sqrt{\frac{\sum_{i=1}^{n} (P_i - T_i)^2}{\sum_{i=1}^{n} (T_i - T)^2}}$$

both the target value (T) and the forecasted value (P) are used.

crop predict

Field Name	Data Type	Width	Description
N	int	15	Nitrogen level in soil
P	int	15	Phosphorous level in soil
K	int	15	Potassium level in soil
PH	int	25	PH value of soil
RF	int	25	Rainfall value in mm
State	int	25	Location
City	int	25	Area

Table 1. Precision and inaccuracy metrics for all classifiers.

Method	Accuracy (%)	Kappa (0~1)	MAE (0~1)	RMSE (0~1)	RRSE (%)	RAE (%)
Bayes Net	98.59	0.998	0.0022	0.023	9.64	1.33
Naïve Bayes Classifier	99.66	0.989	0.0016	0.048	8.73	1.11
Regression	98.48	0.994	0.0056	0.028	23.44	11.77
Decision Table	89.60	0.884	0.0597	0.237	67.41	65.14
Random Forest	98.46	0.996	0.0045	0.190	11.85	3.88
Random Tree	98.44	0.985	0.0664	0.066	19.33	1.39

In cases when particular crop kinds are not known or readily recognizable, this experiment seeks to see whether the algorithm may correctly forecast more generic categories of crops. This could be useful for scenarios where general agricultural features are being researched. Table 2 shows that in order to get good results with machine learning algorithms for crop recognition, it is important to use the right characteristics. Future agricultural data analysis may use our discovered collection of attributes—which includes moisture, temperature, pH, and rainfall—as a guide to pick important features for crop determination.

Thanks to its intuitive design, this online platform makes state-of-the-art AI technology accessible and usable for farmers of all skill levels. The project's goal is to build a sustainable agricultural community by combining data-driven accuracy with resource efficiency. The "Crop prediction and Fertilizer Recommendation with Machine Learning" project aspires to a sustainable agricultural future that guarantees food security, economic viability, with ecological harmony via its comprehensive approach to sustainable precision farming.

To get around the problems with the present system, the proposed solution involves building a comprehensive web platform using Crop Recommendation. The platform utilizes state-of-the-art machine learning technology to offer data-driven insights to farmers, assisting them in making better decisions about crop selection along with nutrient management. To overcome the problems with the current system, farmers are turning to precision farming, which improves agricultural output while decreasing their environmental effect and making better use of available resources.

ADVANTAGES

- Data-driven algorithms making intelligent judgments.
- Personalized advice for effective management of resources.
- Motivating farmers to investigate alternative crop choices.
- Minimized crop failures leading to enhanced yields.
- Environmental protection, resource conservation, and sustainable agricultural methods.

VI. CONCLUSION

Finally, this initiative unveils a revolutionary online platform that equips farmers with cutting-edge technology and insights derived from data. Agriculture can be more productive and environmentally friendly with the help of crop prediction apps that help farmers choose the best crops and make the most efficient use of their resources. To secure food safety and prosperity for future generations, the initiative aims to promote efficient agricultural techniques, reduce environmental impact, and enhance economic viability. It envisions a better future for agriculture as a whole. The project's goal is to make conventional farming more sustainable, efficient, and ecologically aware by using this novel strategy. There is a lot of room for growth in the project, and adding a Yield Prediction component is one exciting step in that direction. In the long run, this project aims to build a powerful Yield Prediction software that uses data to predict harvests. This module may help farmers make better choices, allocate resources more efficiently, and plan their agricultural operations by using weather patterns, soil information, crop-specific traits, and past crop performance.

An all-inclusive system for precision farming will be created when the Yield Prediction component is integrated with the current Crop Recommendation along with Fertilizer Recommendation modules. Accurate yield estimates allow farmers to achieve improved productivity and profitability by making educated decisions regarding crop selection, determining appropriate planting times, and adjusting farming techniques.

An extensive database of past crop yields, meteorological records, and soil data is necessary for a successful implementation of the Yield Prediction module. In order to analyze this data and make accurate yield estimates, statistical models and machine learning algorithms will be crucial. The "Sustainable Precision Farming Techniques with Deep Learning" project will be taken to the next level with the inclusion of the Yield Prediction module. This will provide farmers with essential insights to optimize crop yields, eliminate uncertainties, and create a more resilient and successful agricultural industry. With the rapid advancement of technology, this initiative might completely transform agricultural methods, guaranteeing a future free from food insecurity.

REFERENCES

1. Sharp, Jeff S., and Molly B. Smith. "Social capital and farming at the rural–urban interface: the importance of nonfarmer and farmer relat ions." Agricultural systems 76.3 (2003): 913-927.
2. Shah, Farooq, and Wei Wu. "Soil and crop management strategies to ensure higher crop product ivity within sustainable environments." Sustainability 11.5 (2019): 1485.
3. Priyadharshini, A., et al. "Intelligent crop recommendat ion system using machine learning." 2021 5th international conference on computing methodologies and communication (ICCMC). IEEE, 2021.
4. Capraro, Flavio, et al. "Neural network-based irrigat ion control for precision agriculture." 2008 IEEE International Conference on Networking, Sensing and Control. IEEE, 2008.
5. Rajak, Rohit Kumar, et al. "Crop recommendation system to maximize crop yield using machine learning technique." International Research Journal of Engineering and Technology 4.12 (2017): 950-953.
6. Reddy, D. Anantha, BhagyashriDadore, and AartiWatekar. "Crop recommendation system to maximize crop yield in ramtek region using machine learning." International Journal of Scientific Research in Science and Technology 6.1 (2019): 485-489.
7. Dighe, Deepti, et al. "Survey of crop recommendation systems." IRJET 5 (2018): 476-481.
8. Sardogan, Melike, AdemTuncer, and YunusOzen. "Plant leaf disease detect ion and classification based on CNN with LVQ algorithm." 2018 3rd international conference on computer science and engineering (UBMK). IEEE, 2018.
9. Amara, Jihen, BassemBouaziz, and AlsayedAlgergawy. "A deep learning-based approach for banana leaf diseases classificat ion." Datenbanksystemefür Business, Technologie und Web (BTW 2017)-Authorized licensed use limited to: Zhejiang University. Downloaded on July 17,2023 at 13:30:04 UTC from IEEE Xplore. Restrictions apply. Workshopband(2017).
10. Pudumalar, S., E. Ramanujam, R. HarineRajashree, C. Kavya, T. Kiruthika, and J. Nisha. "Crop recommendat ion system for precision agriculture." In 2016 Eighth International Conference on Advanced Computing (ICoAC), pp. 32-36. IEEE, 2017.