CROP AND FERTILIZER RECOMMENDATION USING AI

[1] SRI RAKSHITHA A K , [2] SHIVA PRANAV S
[1] Final Year, Department of Electronics and Instrumentation Engineering, Bannari Amman Institute of Technology, Sathyamangalam, Erode, Tamil Nadu - 638401
[2] Final Year, Department of Computer Technology, Bannari Amman Institute of Technology, Sathyamangalam, Erode, Tamil Nadu - 638401

ABSTRACT

Instead of cultivating different types in different seasons, farmers produce the same crop every season. They also apply more fertilizer without knowing its exact composition or dosage. In order to describe the best crop to produce and the fertilizer to seed based on soil and weather circumstances, we developed a recommendation model based on machine learning. Therefore, using our technique allows farmers to grow fresh crops in various seasons while making a better profit and preventing soil contamination. Analyze the different factors that are associated, such as the location and the pH level used to calculate the soil's alkalinity. The use of third-party apps like APIs for weather and temperature, soil type, soil nutritional content, degree of rainfall in that zone, and soil composition is combined with the usage of location. Percentages of nutrients like Nitrogen (N), Phosphorous (P), and Potassium (K) are also used. All of these characteristics of the data will be examined while building a model by using several acceptable machine learning techniques on the data. The system includes a model to provide the end user with accurate suggestions about the appropriate fertilizer ratio based on atmospheric and soil data of the field, which enhance crop production and raise farmer income.

Keywords: ***Machine Learning, random-forest, SVM, K-nearest neighbor algorithm, Web application.***

1. INTRODUCTION

India's 1.2 billion citizens work mostly in agriculture, which encompasses 60% of the country's territory and meets their fundamental requirements. Agriculture practices are being modernized nowadays for the benefit of the farmers. Crop yield or output is highly determined by geography, changes in the environment, rainfall (which can be erratic at times), conservation measures, and the use of pesticides. As a result, growers are unable to produce the crop's predicted yield. Currently, a few researchers are using data mining, machine learning, and deep learning approaches to significantly boost output and quality.

By identifying and displaying the consistency and structure of drive data, machine learning enhances unit execution since it may become competent with the machine without characterized computer programming. In this work, a range of machine learning methods, such as Ridge Regression, Gradient Boosting Regression, Random Forest Regression, Decision Tree Regression, Polynomial Regression, and Linear Regression, have been utilized to forecast agricultural yields while optimizing for a range of crops and varied state crop yield datasets.

The system's recommendations for the finest crop for a precise plot of land. based on soil composition and climatic variables such rainfall, temperature, humidity, and pH. These are gathered from the weather forecast, the government website, and V C Farm Mandya. The system uses sensors to collect the necessary data, such as pH, humidity, and temperature, from farmers or other sources.

Support Vector Machine (SVM) and Decision Tree are two machine learning prediction models that use all of the input data in order to identify patterns in the data and then analyze it in conformance with the input parameters. The method suggests a crop for the farmer as well as the quantity of nutrients that should be added for the anticipated crop. The system also depicts the expected yield in q/acre, the volume of seed essential for cultivating in kg/acre, and the crop's price at the moment.

2. LITERATURE SURVEY

Given the variety of research that has been conducted in the subject of agriculture with the use of machine learning, it is considered as a fresh field in this industry. Researchers from all around the world have created and analyzed diverse ideologies in the fields of agriculture and allied disciplines.

[1] CH. Vishnu VardhanChowdary, Dr.K.Venkataramana, For better and excellent tomato crop production, an ID3 algorithm has been developed. It functions on the PHP framework and uses CSV datasets. The various characteristics that this research uses are temperatures, area, humidity, and tomato crop yield.

[2] R. Sujatha and P. Isakki , use data mining methods to make predictions. This model utilized a variety of inputs, including crop identity, land area, soil type, pH value, seed type, and water. It also forecasts plant growth and illnesses, enabling users to select the best crop based on climatic data as well as required inputs.

[3] N. Gandhi, L. J. Armstrong, O. Petkar and A. K. Tripathy, SVM was suggested for predicting rice crop yield. The dataset employed in this method includes a variety of factors, including location, temperature, precipitation, and manufacturing. The classifier used on this dataset is sequential minimal optimization. To create the set of rules on the existing dataset, they prepared the dataset using the Weka tool. SVM algorithm results were developed in Python.

[4] S. Veenadhari, B. Misra and C. Singh , have created the Crop Recommender website, an interactive tool for determining how the climate affects crop productivity. Decision trees and rules have been created based on the c4.5 algorithm. It illustrates how many climate factors influence crop development. Environmental characteristics from the relevant years, such as rainfall and temperature, were studied. The zones under the selected crop determined the options.

[5] Jun Wu, Anastasiya kOlesnikova, Chi- Hwa Song, Won Don Lee [6], a selection tree that is suitable for classifying all varieties of farming records has been proposed. A decision tree classifier was suggested for agricultural data. It makes use of fresh information and may speak to any file. Datasets for soybeans and horse-colic are validated using a 10-fold cross validation approach.

[6] Kiran Mai,C Murali Krishna, I.V, A.VenugopalReddy , In their paper, the article described how data mining is combined with other farming data, such as meteorological conditions, and how the use of pesticides is helpful for ease out of the use of pesticides. Contiguous properties in the agricultural industry were represented using current information.

[7] Verheyen, K., Adrianens, M. Hermy and S.Deckers [8], analyzed the study's use of statistical mining techniques, which are frequently employed to view soil characteristics. K-Means are used in conjunction with GPS-based technology to divide soils into sections.

3. PROPOSED WORK

Based on soil composition and environmental variables like temperature, humidity, soil PH, and rainfall, the proposed method will suggest the best crop and fertilizer for a given plot of land.

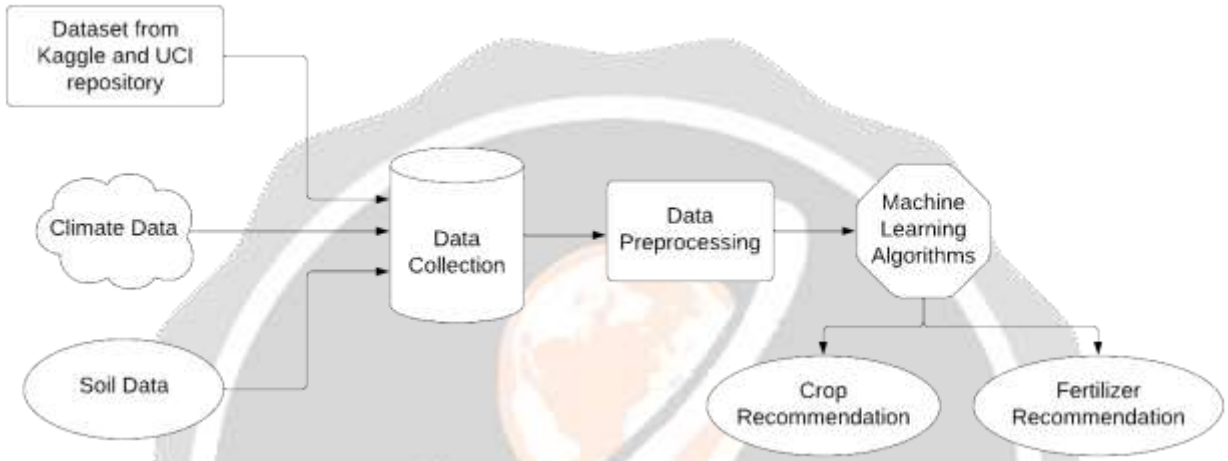

Figure 1: Workflow diagram

The suggested system's architecture is made up of a number of blocks, as indicated in fig(1).

3.1 Data Collection

The most effective way to gather and measure data from various sites, including Kaggle and the UCI machine learning repository, is through data collection. in order to obtain a rough dataset for the system. These characteristics must be present in this dataset. The following variables will be taken into consideration for crop prediction: I Soil PH ii) Temperature iii) Humidity iv) Rainfall v) NPK levels.
Crop Suggestions With the kaggle dataset link provided, you can get the dataset as follows:
https://www.kaggle.com/manojkumardp/crop-recommendation-dataset,

The dataset for fertilizer recommendations can be retrieved similarly as:
https://www.kaggle.com/gdabhishek/fertilizer-predictionare used as training and testing datasets.

3.2 Data-Preprocessing

Datasets were gathered from a variety of sources. Dataset preparation is required before model training. Reading the obtained dataset is the first stage of data preprocessing, which then moves on to data cleaning. As data is cleaned, some redundant attributes are removed from the datasets so that crop predictions can be made. In order to improve accuracy, we must eliminate undesirable attributes and datasets with some missing values. These missing values must also be eliminated or filled with unwanted nan values. Then specify the model's target. The dataset will be divided into a training set and a test set using the Sklearn package after data cleaning.

3.3 AI Algorithms for recommendation

Based on taught data, machine learning predictive algorithms have highly optimized estimations of expected outcomes. Using data, statistical algorithms, and techniques from machine learning, predictive analytics determines the likelihood of future outcomes based on historical data. Beyond simply understanding what has happened, the goal is to provide the most accurate prediction of what will occur in the future. We employed supervised machine learning with the subcategories of classification and regression in our system. For our system, a classification method will work well.

Machine learning algorithms like XGBoost, Random Forest, and KNN are used to forecast the best crop, while SVM and Random Forest are used to predict the best fertilizer.

3.3.1 Crop Recommendation

Input characteristics like N, P, K temperature, humidity, and rainfall are used to predict the specific crop that will be cultivated. The loading of external crop datasets is the first step in the crop forecast process. After the dataset has been read, different phases of pre-processing will be carried out, as mentioned in the section on data pre-processing. After pre - processing the data, create a training dataset and train the models using KNN and Random Forest classifier. We take into account a number of variables, including temperature, humidity, soil PH, and anticipated rainfall, when predicting the crop. These are the system input parameters that can be manually inputted or obtained through sensors. We'll append a list with the predicted rainfall totals and input parameter values.

The following table lists a Summarized Crop forecasting models and its input parameters:

N	P	K	Temparature	Humidity	Rainfall	Crop
90	42	43	20.87	82.00	202	rice
85	58	41	21.77	80.31	226	rice
60	55	44	23.00	82.32	263	rice
74	35	40	26.49	80.15	242	rice

Table-1: Input Parameters Values with Output for Crop Recommendation

Research on Crop Data below are some instances of agricultural analyses that were conducted using univariate and bivariate data sets.

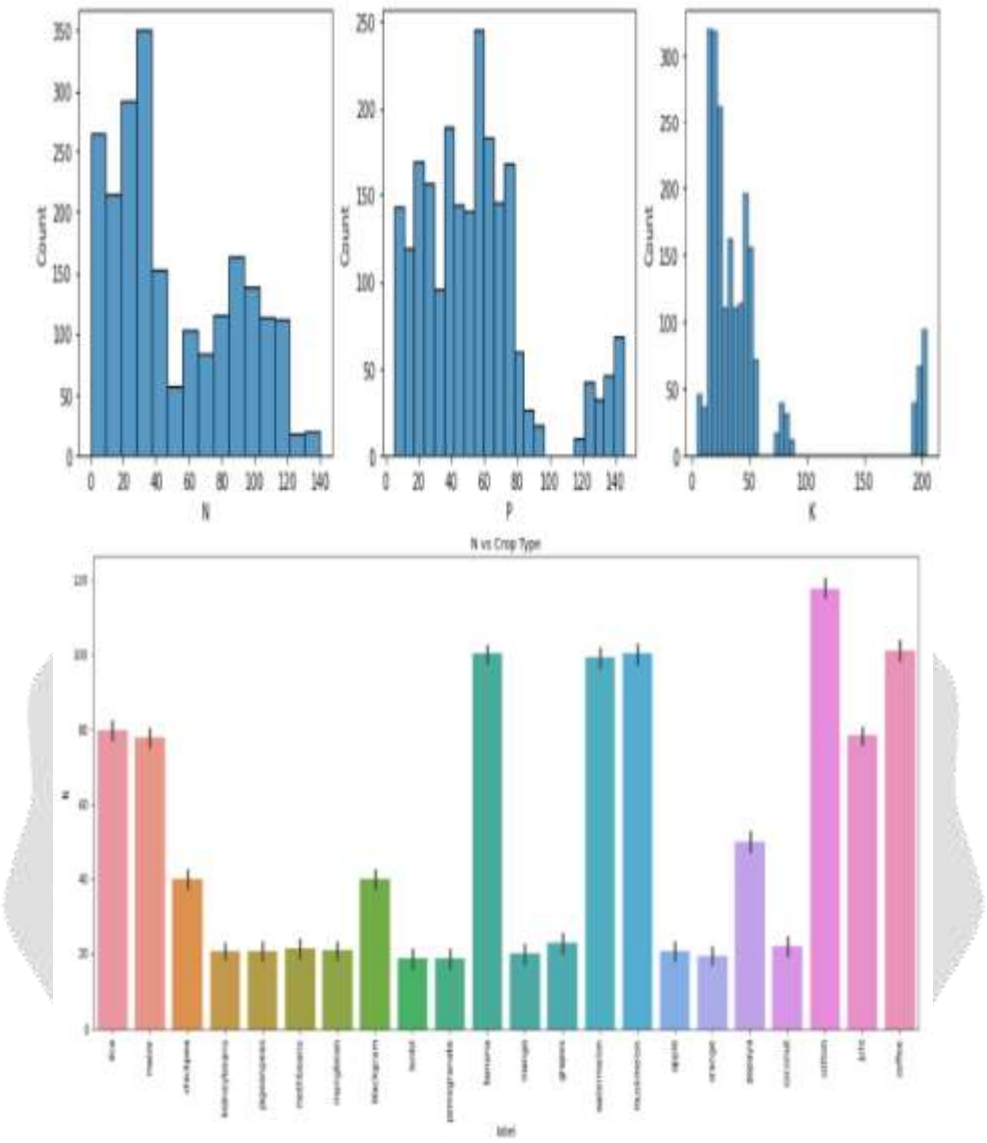

Figure 2 : Univariate and bivariate analysis for order to ensure the sustainability parameters N, P, and K

3.3.2 Fertilizer Recommendation

We use input data like N, P, K temperature, humidity, wetness, soil type, and crop to be cultivated in order to predict the specific fertilizer to be applied. Fertilizer prediction method begins with loading the external fertilizers datasets. After the dataset has been read, different phases of pre-processing will be carried out, as mentioned in the section on data pre-processing. Train the models using SVM and Random Forest classifier into a training dataset after data pre-processing. We take into account numerous factors, like temperature, humidity, soil PH, and the anticipated crop to be cultivated, while predicting the fertilizers. These are the input parameters for a system that can be manually inputted or obtained through sensors. Values for the input parameters and the predicted crop will be appended to a list.

The following table lists a summarized fertilizer prediction model and its input parameters:

N	P	K	Temparature	Humidity	Rainfall	Crop
90	42	43	20.87	82.00	202	rice
85	58	41	21.77	80.31	226	rice
60	55	44	23.00	82.32	263	rice
74	35	40	26.49	80.15	242	rice

Table-2: Input Parameters Values with Output for Fertilizer Recommendation

Analysis of Fertilizer Data here are some examples of fertilizer analysis using a dataset that was run through univariate and bivariate analysis.

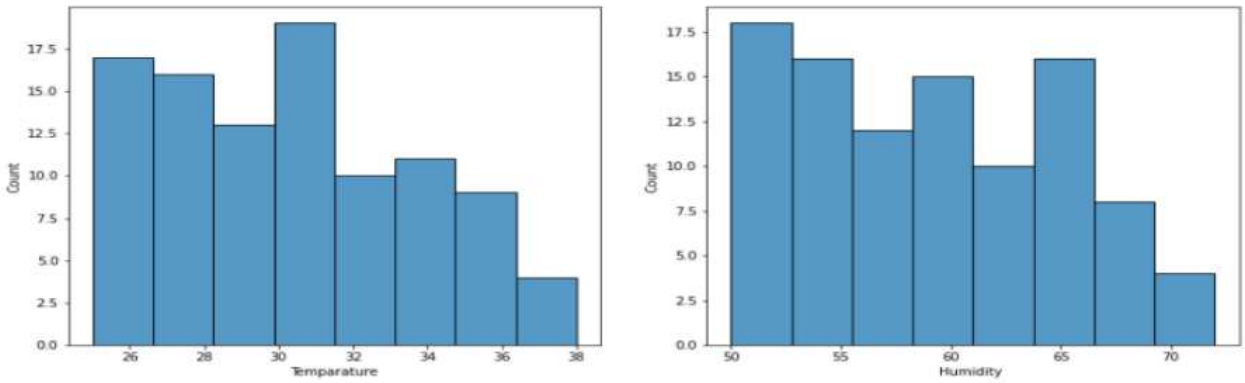

Figure 3: Study of N, P, and K parameters for fertilizer prediction using univariate and bivariate methods

4. SUPPORT VECTOR MACHINE

For classification and prediction, Support Vector Machines (SVM) are a machine learning technique. With this method, each data point is represented as a point in n-dimensional spaces, and each result is expressed by a required in order that is plotted as a hyper-plane.

To categorize provided data, an SVM constructs hyperplanes with the biggest margins in a high-dimensional space.

The distance between the two classes' closest data points is represented by the margin between them.
STEP 1: Choose the feature sets from the various data classes.
STEP 2: Determine the spots where each class of feature and plot intersect. Repeat for all the data features.
STEP 3: Delete the data from all classes as well as the features that intersect them.
STEP 4: involves plotting the hyper planes for the remaining spots.
STEP 5: Determine the hyperplane distances for various classes of objects.
STEP 6: is to choose the hyperplane that is consistent across all data classes.

5. RESULTS AND DISCUSSION

By taking into account factors like annual rainfall, temperature, humidity, and soil pH, the proposed approach suggests the optimal crop for a certain plot of land. While some of these parameters—like yearly rainfall—are predicted by the system using data from the prior year and the SVM algorithm, others require user input. The system uses NPK values from the input area to present the recommended crop's required NPK, as well as a suitable crop, the number of seeds needed per acre, the market price, and an approximation of the crop's yield.

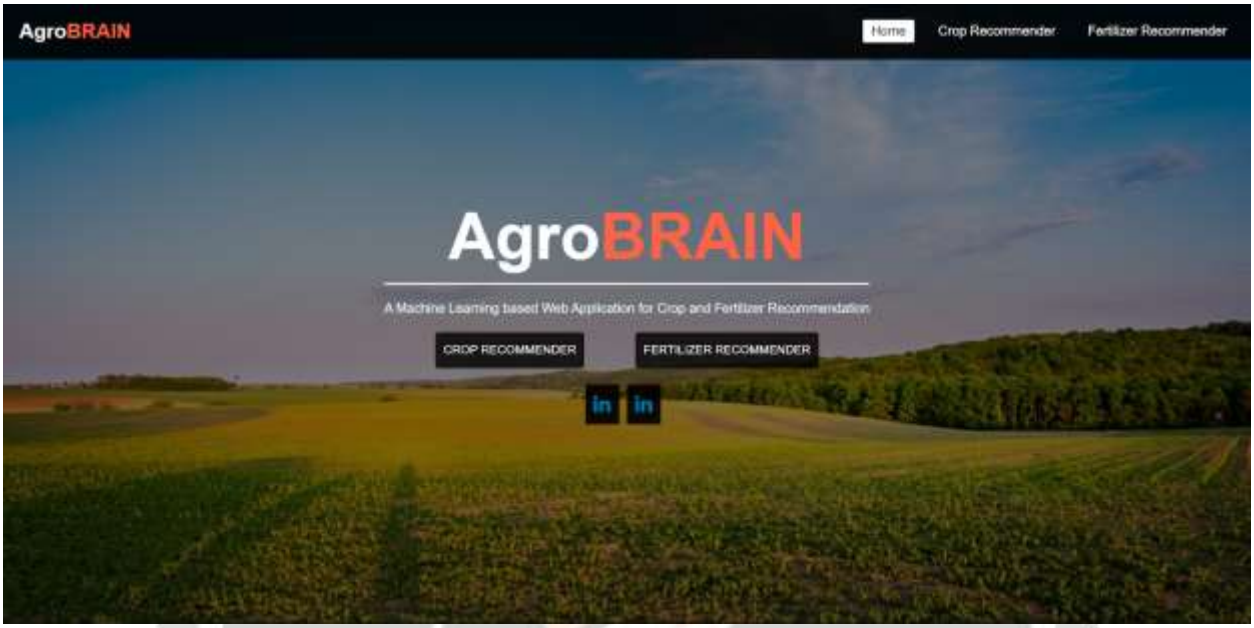

Figure 4: Home page of the web application

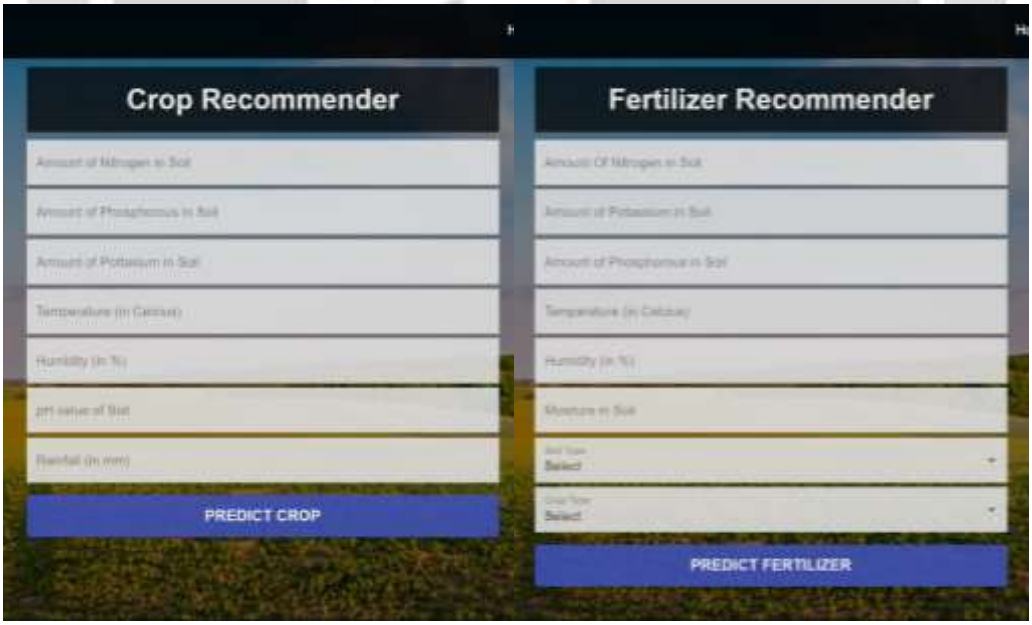

Figure 5,6: Crop and Fertilizer Recommender page

Figure 7,8 Recommended crop and fertilizer for user input

6. CONCLUSION

As our farmers are now not employing technology and analysis properly, there is a danger that they will choose the wrong crop to cultivate, which will lower their income. To lessen these kinds of losses, we have created a farmer-friendly system with a graphical user interface (GUI) that will forecast which crop would be the greatest fit for a specific plot of land. This system will also provide details on the nutrients that must be added, the seeds that must be used for cultivation, the anticipated yield, and the market price. So, this encourages farmers to choose the best crop for cultivation so that the agricultural sector can be enhanced through creative thinking.

7. REFERENCES

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

[2]. Alakeel, T., & Elhoseny, M. (2021). A comprehensive survey on precision agriculture using artificial intelligence. Computers and Electronics in Agriculture, 180, 105956.

[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]. Tewari, A., & Mishra, S. K. (2020). Application of artificial intelligence in agriculture: A comprehensive review. International Journal of Computer Applications, 179(40), 11-19.

[5]. Pabon, J. D., Arango, J. A., & Rodriguez, R. E. (2021). An approach for fertilization recommendation using machine learning algorithms. In 2021 IEEE 23rd International Conference on High Performance Computing and Communications; IEEE 19th International Conference on Smart City; IEEE 7th International Conference on Data Science and Systems (HPCC/SmartCity/DSS) (pp. 1079-1086). IEEE.