GrowFarm – Crop, Fertilizer and Disease Prediction usingMachine Learning

Suresh Singh Rajpurohit[1], Sowmen Parui[2], Rakshita Bhole[3]

[1] Assistant Professor, Dept. of Information Technology, Thakur College of Science and Commerce,Mumbai, India.
[2] Undergraduate Student, Dept. of Data Science, Thakur College of Science and Commerce, Mumbai,India.
[3] Undergraduate Student, Dept. of Data Science, Thakur College of Science and Commerce, Mumbai,India.

Abstract - The farming industry is extremelyimportant in India for economic growth and employment creation. Agriculture employs around 48% of the inhabitants in India. It gives locals the opportunity to work and contribute to the growth of a country like India on a large scale, as well as strengthens the economy, as agriculture is the backbone of India's developing economy. Farmers have always followed historical agricultural techniques and traditions. But, a single farmer cannot beexpected to consider all of the numerouselements that influence crop development. A single erroneous judgment made by the farmer might have unfavorable consequences. The project's goal is to help farmers determine the quality of their soil and analyze its many properties, as well as to propose crops and fertilizers based on the results of a machine-learning approach. The system utilizes multiple Categorization strategies to increase the performance of the Crop Recommendations System and Fertilizer Recommendation System. As a result, the strategy assists novice farmers in gathering information. This project takes soil and ph dataas inputs and creates a website to anticipate which crops are most suited to the soil and which fertilizers may be used to treat illnesses discovered in plants.

Keywords - Crop Recommendation, Fertilizer Recommendation, Plant Disease Detection, Smart Farming.

INTRODUCTION

Agriculture is a major industry in India, employing the majority of the population. As the world's population expands, so do agricultural challenges. It is one of the most important vocations for human survival. We have noted that the environment is always changing, which is damaging crops and drivingfarmers into debt and suicide. The majority of farmers face the problem of planting the wrong crop for their land based on a traditional or non-scientific approach. This is a daunting task for a country like India, where agriculture feeds over 42% of the population. Crop selection mistakes lead to reduced yield and profit. As a result, farmers are relocating to cities for jobs, attempting suicide, giving up farming, leasing land to industrialists, or utilizing it for non-agricultural purposes.

Because resources are limited, it is vital that they be used carefully and efficiently. In this setting, technology is critical since it may help solve problems and prevent resource waste by analyzing and anticipating circumstances. The proposed system is being implemented using machine learning, which is one of the applications of Artificial Intelligence. Crop recommendation will offer the ideal crop for your property based on soil nutrient value and climate in that region.

The ensemble method is used to create a recommendation model that integrates the predictions of various machine learning models to select the optimum crop and fertilizer to utilize based on soil value. One of the most important aspects of a good farming system is disease identification. In general, a farmer monitors disease symptoms in plants that require regular monitoring through eyeobservations. Various types of illnesses, damage plant leaves. Farmers encounter increased challenges in recognizing these illnesses, so we utilize image processing methods that are acceptable and efficient with the help of plant leaf images for disease identification.

Objective and scope of the project

The project's goal is to create a croprecommendation system to create a robust model suitable for predicting crop sustainability in a particular state based on soil type and climatic parameters. It recommends the best crops for the region so that farmers can reduce their losses while farming. This project also suggests which fertilizer to use as well as how to care for plants and crops. To assist people in identifying unfamiliar crops and plants, as well as to detect crop disease. It serves as a one-stop shop for an individual to obtain all basic knowledge on plants and crops,as well as their uses and benefits, in one location. This approach can give a mechanism for crop, fertilizer, and crop disease prediction.

LITERATURE SURVEY

This section summarises the conclusions of multiple articles that have been studied and reviewed. This section contains records that were reviewed prior to and during project development. The documents provided an improved understanding of existing solutions, how methods can be optimized, and how algorithms could be selected based on their performance to get a better result while developing the Project.

Title and Authors	Year of publication	Work Done	TechniquesUsed
Crop Recommendation System for Precision Agriculture	2016	SAD is a crop selection method that improves net yield rate by improving accuracy and classification performance.	Support Vector Machine, Random Forest, Neural Network, REPTree, Bagging, and Bayes
Crop Recommender System Using Machine Learning Approach	2021	API is used to predict weather parameters.ML algorithms are used to compare the results of models. Random Forest is the algorithm used with an accuracy Of 95%.	ANN, Fuzzy Network, decision tree, KNN, Enet, Lasso and Kernel Ridge, ARMA, SARIMA and ARMAX
Improvement of Crop Production Using Recommender Systemby Weather Forecasts	2019	ANNs are usedto predict rainfall events and categorize them, resultingin a hybrid approach with an accuracy percentage of 96%.	ANN, DNN, Layer Recurrent networks, FFBPNs, and Cascaded-Feed Forward-Back Propagation networks, Naïve Bayes
Improving Crop Productivity through a Crop Recommendation System Using Ensembling Technique	2018	Ensembling technique to improve yieldof crops, pre-processingand classifyingthem using thesoil dataset. The accuracy obtained fromthis techniqueis 99.91%.	Naïve-Bayes,k-Nearest, Random Forest, Linear SVM, Majority Voting
A Recommended System for Crop Disease Detection and Yield Prediction Using Machine Learning Approach	2020	A system to detect disease and predict crop yield using machine learning to maximize production on a limited land resource.	ANN, KNN, SVM, Decision Tree Learning, Random Forest, Gradient Boosted Decision Tree, Regularized Greedy Forest
Agricultural Crop Recommendations based on Productivity and Season	2021	The system uses collaborative and content-based filtering to analyze agriculture parameters using machinelearning techniques.	K- means, K-means++, JRip, J48, Naïve Bayes, neural networks, soft computing, SNN model, GIS, ANN, KNN
Efficient Crop Yield Recommendation System Using Machine Learning For Digital Farming	2021	The Supervised Learning Algorithm is used to predict harvests with higher precision and productivity forranchers to settle on the yield to be planted.	Backing Vector Machine Calculation to perform AI, Supervised machine learning algorithm

Table 1: Dataset of Crop Recommendation (First 10 rows)

Nitrogen (N)	Phosphorus (P)	Potassium (K)	Temperature	Humidity	ph	Rainfall	Label
72	51	44	20.87974371	82.00274423	6.502985292	202.9355362	Rice
99	56	44	21.77046169	80.31964408	7.038096361	226.6555374	Rice
69	42	38	23.00445915	82.3207629	7.840207144	263.9642476	Rice
73	44	45	26.49109635	80.15836264	6.980400905	242.8640342	Rice
73	51	41	20.13017482	81.60487287	7.628472891	262.7173405	Rice
63	48	35	23.05804872	83.37011772	7.073453503	251.0549998	Rice
96	38	41	22.70883798	82.63941394	5.70080568	271.3248604	Rice
98	55	44	20.27774362	82.89408619	5.718627178	241.9741949	Rice
65	53	41	24.51588066	83.5352163	6.685346424	230.4462359	Rice
75	58	39	23.22397386	83.03322691	6.336253525	221.2091958	Rice

METHODOLOGY

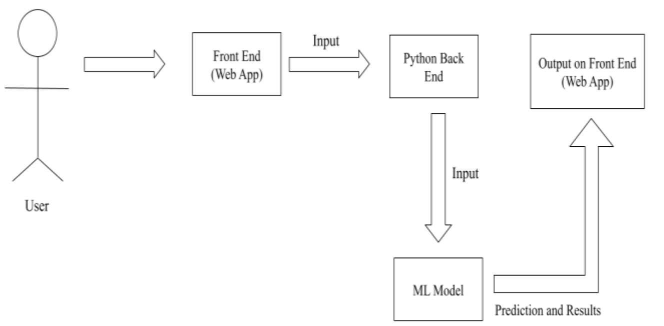

Figure 1: Architecture Diagram

We propose a system with a simple, cost-effective, user-friendly User Interface that is also time efficient. Our proposed approach assists farmers and users in achieving their objectives. This method recommends crops and fertilizers while also forecasting plant problems. In this proposed system, we will collect factors such as nitrogen, phosphorus, potassium, and others and recommend crops or fertilizers using methods such as Decision Tree, Random Forest, Naive Bayes, Support Vector Machine, and Logistic Regression, which will aid in accurate prediction. In addition, in this proposed system, we will take a picture of the plant then the algorithm will predict the disease using the ResNet algorithm. As a result, this approach will make farming easier while also increasing customer satisfaction.

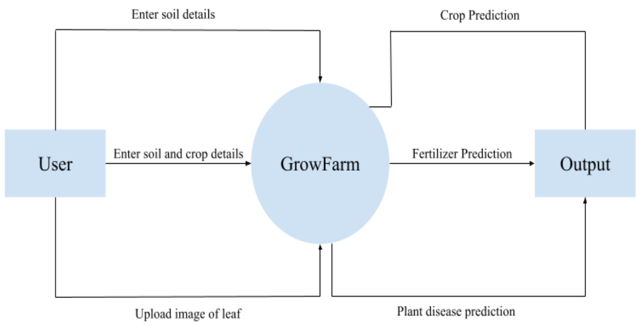

Figure 2: Data Flow Diagram

Step 1:

Loading the dataset, the data which was collected from the study Crop will be used in the system to optimize Crop Production Using Machine Learning Algorithms, which included a crop recommendation dataset. The accuracy of a machine learning model is determined by the quality of the data.

Step 2:

Pre-processing of the input dataset, the most essential or time-consuming task in any machine learning project is data pre-processing. During the pre-processing step, missing values are filled using techniques such as mean, mode, and median, scaling or transforming values in a certain range, cleaning the data, encoding categorical data, and checking for variable correlation so that the accuracy can be increased.

Step 3:

Analyzing exploratory data, before getting hands dirty with model construction, univariate, bivariate, and multivariate analysis are carried out to uncover hidden patterns in the data and to try to interpret the data. A few examples of the univariate analytic plot are PDF and CDF while multivariate plots are pair plot, box plot, and Heatmaps.

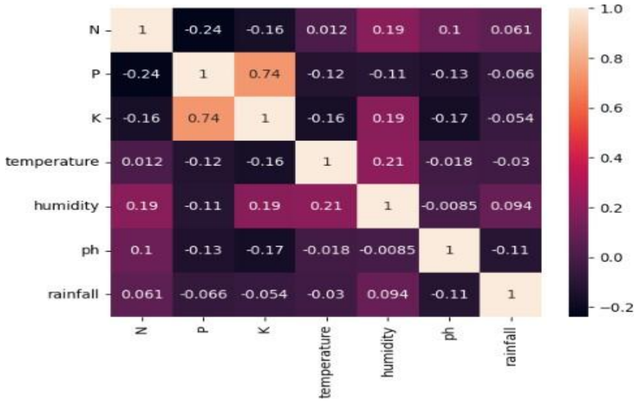

Figure 3: Correlation Metrics on the croprecommendation dataset

Step 4:

Splitting data into training and testing in this step, the pre-processed dataset is divided into training and testing groups based on 80:20 ratios, which indicates that 80% of the data is used for training and 20% for testing on the unseen dataset and cross-validation to discover the best hyper parameter.

Step 5:

Creating a classification model based on the training data, the training dataset is delivered to the individual classifier in this stage, and the model is trained on top of it.

Decision Tree:

A decision tree is a supervised machine-learning technique that can be applied to both classification and regression. A decision tree has a structure similar to a flowchart, with attributes and class labelsdisplayed by a tree.

Logistic Regression:

Logistic regression is any other efficient supervised ML set of guidelines used for binary categorization problems (while the goal is categorical). Logistic regression uses a logistic characteristic mentioned below to model a binary output variable. The primary distinction between linear regression and logistic regression is that logistic regression has a range of zero to one. Furthermore, logistic regression, as opposed to linear regression, no longer requires a linear connection between input and output variables. This is due to the employment of a non-linear log transformation on the chance ratio.

Random Forest:

Random Forest is a machine-learning system that is built on ensembles. Ensemble approaches are a type of method that allows us to mix independent or similar algorithms to construct a powerful model. A random forest is a collection of multiple decision trees that have the highest depth until the nodes can separate with the least variability and bias.

Gaussian Naive Bayes:

Gaussian Naive Bayes is a simple and straightforward machine learning technique.According to the Naive Bayes hypothesis,qualities must be independent of one another. Internally, the Bayes theorem is applied in Binary Classification, which is a statistical based technique. If the characteristics of the dataset have a Gaussian distribution, the model is known as Gaussian Naive Bayes.

Support Vector Machine:

SVM stands for Support Vector Machine algorithm which is a machine learning technique. First, it plots each data element inN Dimension space and then selects the hyper-plane that best segregates the two classes with the greatest margin in the Linear Kernel. It is tough to find the optimal hyper-plane in a Support vector machine.

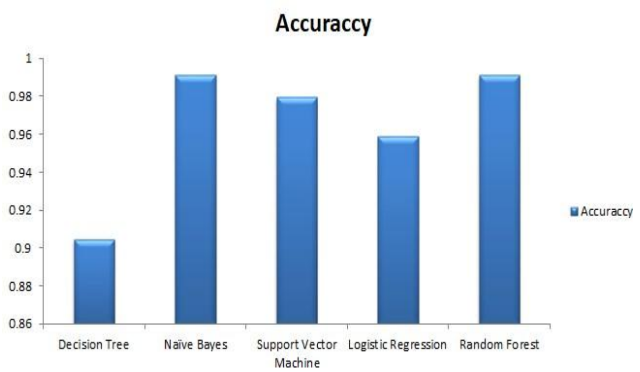

Figure 4: Accuracy comparison chart of differentalgorithms used in the classification model

ResNet :

This network's topology is designed to allow vast amounts of fully connected layers to feature efficiently. Yet, adding a number of deep layers to a network frequently causes output degradation. This is known as the vanishing gradient problem, in which neural networks, while learning via reduced back propagation, rely on gradient descent, descending the loss feature to identify the minimizing weights. Because of the presence of several layers, the repeated multiplication effects inside the gradient get less and smaller, "vanishing," leading to saturation within the community's overall performance or maybe worsening the overall performance. The primary principle of ResNet is the use of leaping connections, also known as shortcut connections or identification connections. These connections are typically formed by jumping over one or two layers, forming shortcuts between those layers. The goal of establishing those shortcut connections was to tackle the major problem of vanishing gradient encountered by deep networks. These identification mappings, for starters, no longer perform anything more than bypass the connections, resulting in the employment of previous layer activation.

Future Work

To improve results and support, the system can be expanded further by adding the following functionality:

The focus of future work will be on upgrading datasets on a regular basis to create reliable forecasts, and the process can be automated without modifying the dataset manually. Linking the system to physical devices to transform it into an IOT device that can check soil components without human intervention and select crops to grow depending on the results. Providing users with real-time crop market rates. Multilingual communication is possible, allowing people from all over the world to use this system.

CONCLUSION

This system developed with machine learning known as Crop Recommendation or Prediction system will assist in recommending the best crop to grow inland, as well as which fertilizer to use, and provide plant disease detection based on images, which will be easily available and used by users in order to make a decision on which crop to grow based on the soil nutritional values and climate in that region. The model proposed in the research can be expanded in the future to include crop and fertilizer recommendations as well as plant disease detection in a mobile app. Consequently, our website will assist farmers in sowing the appropriate seed based on soil requirements and increasing their cropyields in order to boost production and profit in their operations (if any) from suchtechniques.

REFERENCES

1. S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika and J. Nisha, "Crop recommendation system for precision agriculture," 2016 Eighth International Conference on Advanced Computing (ICoAC), 2017, pp. 32-36, doi:10.1109/ICoAC.2017.7951740. IEEE

2. S. M. PANDE, P. K. RAMESH, A. ANMOL, B. R. AISHWARYA, K. ROHILLA and K. SHAURYA, "Crop Recommender System Using Machine Learning Approach," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp.1066-1071, doi: 10.1109/ICCMC51019.2021.9418351. IEEE

3. S. Bangaru Kamatchi, R. Parvathi, Improvement of Crop Production Using Recommender System by Weather Forecasts, Procedia Computer Science, Volume 165, 2019, Pages 724- 732, ISSN 1877-0509. ScienceDirect

4. N. H. Kulkarni, G. N. Srinivasan, B. M. Sagar and N. K. Cauvery, "Improving Crop Productivity Through A Crop Recommendation System Using Ensembling Technique," 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), 2018, pp. 114-119, doi: 10.1109/CSITSS.2018.8768790. IEEE

5. Akulwar, Pooja. (2020). A Recommended System for Crop Disease Detection and Yield Prediction Using Machine Learning Approach. 10.1002/9781119711582.ch8. ResearchGate

6. S. Vaishnavi., M. Shobana., R. Sabitha. and S. Karthik., "Agricultural Crop Recommendations based on Productivity and Season," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021, pp. 883-886, doi: 10.1109/ICACCS51430.2021.9441736. IEEE

7. Dr.. G.Suresh, Dr. A.Senthil Kumar, Dr.S.Lekashri, Dr. R.Manikandan, "Efficient Crop Yield Recommendation System Using Machine Learning For Digital Farming", IJMA, vol. 10, no. 1, pp. 906 - 914, Mar. 2021.

8. Sameer Panat , Acropolis Institute Of Technology And Research , Indore; Siddhant Tripathi, Acropolis Institute Of Technology And Research , Indore; Sarthak Bhaiji, Acropolis Institute Of Technology And Research, Indore; Vandana Kate, Acropolis Institute Of Technology And Research , Indore; Kavita Namdev, Acropolis Institute Of Technology And Research , Indore. IJSRD.

9. S. S. Kale and P. S. Patil, "A Machine Learning Approach to Predict Crop Yield and Success Rate," 2019 IEEE Pune Section International Conference (PuneCon), Pune, India, 2019, pp. 1-5, doi: 10.1109/PuneCon46936.2019.9105741.

10. R. Kumar, M. P. Singh, P. Kumar and J. P. Singh, "Crop Selection Method to maximize crop yield rate using machine learning technique," 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), Avadi, India, 2015, pp. 138-145, doi: 10.1109/ICSTM.2015.7225403.

11. V. Kumar, H. Arora, Harsh and J. Sisodia, "ResNet-based approach for Detection and Classification of Plant Leaf Diseases," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 495-502, doi: 10.1109/ICESC48915.2020.9155585.

12. Ramkumar, M., Catharin, S. S.,Ramachandran, V., & Sakthikumar, A. (2021). Cercospora Identification in Spinach Leaves Through Resnet-50 Based Image Processing. Journal of Physics, 1717(1), 012046, doi: 10.1088/1742-6596/1717/1/012046.