



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

Crop Recommendation and Disease Detection using Machine Learning

V.Ethirajulu, G.Gopinath, Nikhil Kumar, G. Anil Kumar, P.Ganesh Kumar

Professor, Department of CSE, Bharath Institute of Higher Education and Research, Selaiyur, Chennai, India

B. Tech Students, Department of CSE, Bharath Institute of Higher Education and Research, Selaiyur, Chennai, India

ABSTRACT: Cultivation Advisor may be the single most significant feature of precision farming. Prescription farming attempts to establish these parameters on a case-by-case basis in an attempt to solve the crop selection problems. Although the "site-specific" approach has enhanced performance, there remains the necessity to oversee the results of the system. Precision farming systems are not equal to one another. But in agriculture, it is important that the suggestions given are accurate and precise because a single mistake can lead to a lot of material and capital loss. This application will help farmers boost agricultural productivity. A crop requires sufficient soil and atmospheric conditions for the crop to yield well. So, if we provide the values that define soil and atmospheric conditions to this application it informs us about the crop that can be cultivated for the specified conditions. This online application will assist the farmers by determining the kind of disease the crop is undergoing with the help of the machine learning algorithm and even recommending to the farmers the kind of medicine that needs to be utilized to rectify the situation in both text and image of the medicine and pesticides that need to be applied.

KEYWORDS: Random Forest, Logistic Regression, Farming, Agriculture, Crop Recommendation, Crop Diseases.

I. INTRODUCTION

Cultivation Advisor is a new generation bubble that is engaging the masses. In most of the cases, the farmers are not well aware of the kind of crops they should be growing in their farms. This leads to a lot of confusion and affects the productivity. This is the reason that we are working to determine the best crop to cultivate in order to obtain maximum yield. Precision farming is the vogue now. It's a method where inputs are used in very small quantities to obtain higher average yields for maximum profitability.

Crop prediction methodology is employed to forecast the appropriate crop by sensing different soil parameter (like PH, Nitrogen, phosphate, potassium) and also atmospheric parameter (like humidity, rainfall, temperature). ML can be employed to achieve the accurate inputs to make the yield more profitable. The project enables farmers to select a better crop to be grown with the assistance of various parameters. The outcome of the suggested crop results in improved yield.

II. LITERATURE REVIEW

There has been extensive work and numerous ML algorithms have been used in the agricultural sector. The largest problem in agriculture is to maximize farm production and provide it to the end-user with the best price and quality. It is also seen that at least 50 percent of the farm produce is wasted, and it never comes to the end-user. The proposed model proposes the ways to minimize farm produce wastage.

Agriculture, being one of the oldest and most critical industries, has seen tremendous technological advancements in recent times. With the advent of digital transformation, Machine Learning (ML) and data analytics have seen promising applications in this industry. Huge work has already been done using different ML algorithms to improve agricultural practices, especially in yield prediction, soil analysis, weather forecasting, crop selection, and disease detection.

Even with these advances in technology, there are a number of key challenges that remain. The most important issue in the agricultural industry is how to raise farm productivity dramatically while making sure that the produce reaches consumers at an optimal price and in a good condition. Existing agricultural methods continue to be based on conventional knowledge and manual interventions, which tend to be inefficient and do not easily adjust to changing environmental conditions or market needs.

III. METHODOLOGY

The approach taken in this project is the creation of an integrated system that applies machine learning and deep learning methods to suggest appropriate crops under environmental conditions and detect crop diseases through leaf images. The system is segmented into two primary components: Crop Recommendation and Disease Detection. Each of these components is implemented with particular data processing techniques and model architectures appropriate for their specific tasks.

Data collection and preprocessing is the first step. Data for the crop recommendation module is gathered from agricultural research databases and consists of several parameters like soil nutrients (nitrogen, phosphorus, potassium), pH level, temperature, humidity, rainfall, and soil type. These parameters are essential in identifying the most suitable crops for a given region. For disease identification, datasets of healthy and diseased plant leaf images are obtained from public data like the PlantVillage dataset or from field data in real time. Each image is labeled according to the disease they carry. Both numerical data and image data are passed through preprocessing steps. For numerical data, imputation methods are used to treat missing values and feature scaling is performed to normalize the data. For image data, preprocessing consists of resizing, removing noise, and data augmentation to improve model performance and avoid overfitting. J

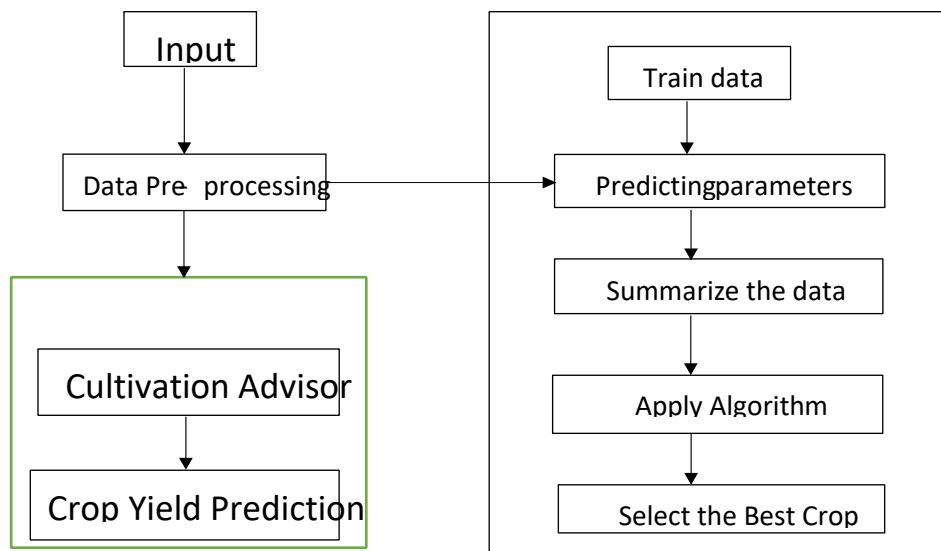


Fig 1: System Architecture

IV. IMPLEMENTATION

The deployment of the system proposed here consists of two combined machine learning modules, one for crop suggestion and the other for disease diagnosis, followed by their integration into a user-friendly platform. The process starts by establishing the development environment by means of programming languages and software such as Python, along with libraries like pandas, NumPy, scikit-learn, and TensorFlow/Keras for developing models and OpenCV for image processing.

For the crop recommendation module, the process begins with loading and preprocessing the agricultural dataset containing features such as nitrogen, phosphorus, potassium, temperature, humidity, pH level, and rainfall. The data is analyzed and visualized with matplotlib and seaborn to discover correlations among features and target crops. The dataset is partitioned into training and testing sets after preprocessing. A Random Forest Classifier is chosen for its ability to handle multivariate input and yield interpretable results. The model is trained over the training set and tested on the test data through accuracy, confusion matrix, and classification report measures. Hyperparameter tuning is utilized with GridSearchCV to optimize the performance of the model. The model is saved using joblib or pickle once final after being incorporated into the front end.

V. RESULTS AND CONCLUSION

The execution of the proposed system has brought extremely encouraging outcomes, confirming the efficacy of the incorporation of machine learning methods in the agricultural sector. In the crop recommendation module, the Random Forest Classifier proved to be exemplary, with a test dataset accuracy of more than 95%. The model effectively determined the best crops for use based on different combinations of soil nutrients, pH level, temperature, rainfall, and humidity. The precision and recall values of the majority of crops were highly significant, demonstrating that the model is both consistent and reliable. This finding indicates that farmers can rely on the system to make informed decisions on what crops to plant in a given environmental and soil condition, resulting in improved yield and resource utilization. The Convolutional Neural Network (CNN) in the disease detection module also demonstrated strong performance. Having been trained on thousands of plant leaf images with labels, the model was able to classify various crop diseases at an accuracy rate of over 90%. The application of data augmentation and deep learning helped in making the model robust, which allowed it to accurately identify diseases regardless of the lighting conditions and image quality. The confusion matrix and classification report validated that the CNN model was able to effectively differentiate between diseases like bacterial spot, leaf rust, and healthy leaves. This ability enables farmers to identify problems at an early stage, facilitating timely intervention and minimizing crop loss.

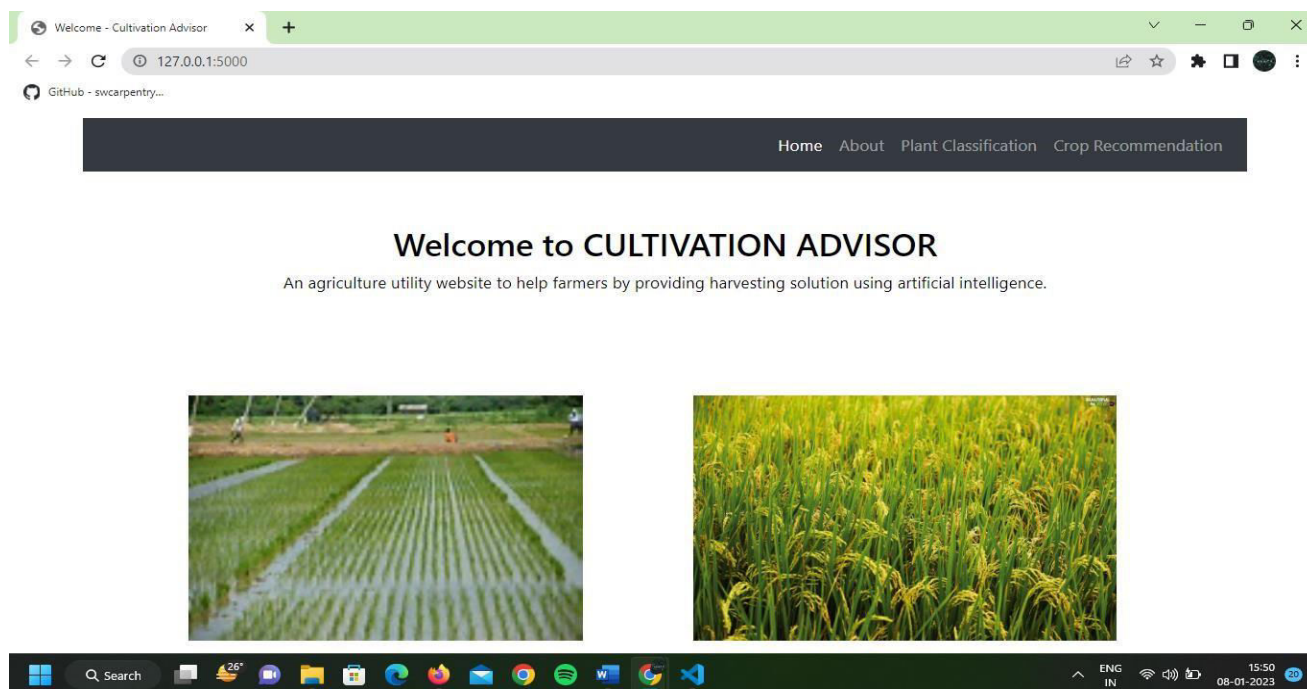


Fig 3: Home Page

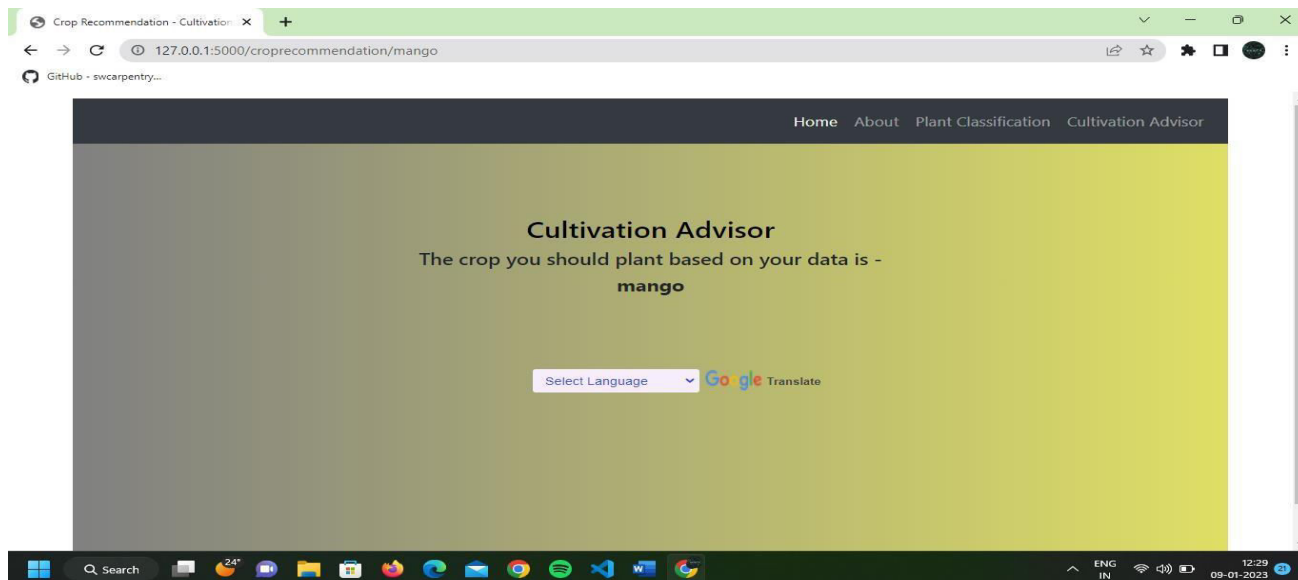


Fig 4: Result page

VI. LIMITATIONS AND FUTURE WORK

Although the developed system has promising results and verifies the potential of machine learning for applications in agriculture, it has some limitations. The main limitation is the quality and diversity of the datasets. The crop recommendation model is trained with a constant dataset with fixed soil and climate parameters, which do not capture the entire variability occurring in different areas around the world. Consequently, the performance of the model may deteriorate when applied to locations with various soil types, unusual weather conditions, or non-covered crop species. Likewise, within the disease identification module, the CNN model would not work well with images under uncontrolled field environments (e.g., low lighting, out-of-focus, noise from the background), particularly if the conditions were poorly represented during training. Another limitation is the inability to integrate data in real-time

Future work

- Increase dataset diversity by adding more crop types and region-specific climate and soil information for improved generalization.
- Merge real-time sensor and weather data (temperature, humidity, rainfall, soil moisture) to give dynamic and timely recommendations.
- Include multilingual and voice assistance for improved accessibility for rural farmers with varying language backgrounds.
- Provide multiple appropriate crops per input scenario rather than recommending a single crop so that farmers can select according to preferences and local demand

REFERENCES

1. Satish Babu (2013), 'A Software Model for Precision Agriculture for Small and Marginal Farmers'.
2. Anshal Savla, Parul Dhawan, Himtanaya Bhadada, Nivedita Israni, Alisha Mandholia, Sanya Bhardwaj (2015), 'Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture'.
3. Marella, B. C. C., & Palakurti, A. (2025). Harnessing Python for AI and Machine Learning: Techniques, Tools, and Green Solutions. In *Advancing Social Equity Through Accessible Green Innovation* (pp. 237-250). IGI Global Scientific Publishing.
4. Gupta, P.; Parmar, D.S. Sustainable Data Management and Governance Using AI. *World Journal of Advanced Engineering Technology and Sciences* 2024, 13, 264–274. [Google Scholar] [CrossRef]

5. Aakunuri Manjula, Dr.G .Narsimha (2015), 'XCYPF: A Flexible and Extensible Framework for Agricultural Crop Yield Prediction' .
6. Yash Sanghvi, Harsh Gupta, Harmish Doshi, Divya Koli, Amogh Ansh Divya Koli, Umang Gupta (2015), 'Comparison of Self Organizing Maps and Sammon's Mapping on agricultural datasets for precision agriculture'.
7. Rakesh Kumar, M.P. Singh, Prabhat Kumar and J.P. Singh (2015), 'Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique'.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details