IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

OPTIMAL CROP RECOMMENDATION FROM CROP YIELD PREDICTION WITH SOIL IMAGE RECOGNITION

[1] Dr. D. Uma Devi,[2]Dr.K.Swapna,[2]P. Sai Likhith, [4]P. Sandeep
[1]Associate Professor (Hod),[2]Assistant Professor(Ad-hoc),[3]Student, [4]Student
[1,3,4] Computer Science and Engineering,[2]Information Technology and Computer Applications
[1,3,4]Gayatri Vidya Parishad College of Engineering (Autonomous), Visakhapatnam, India,[2]Andhra University College of Engineering),Visakhapatnam, India

Abstract: In our comprehensive agricultural project, we tackle the pivotal challenge of crop yield prediction by leveraging advanced machine learning methodologies. Our innovative approach integrates soil image recognition and a range of pertinent parameters to forecast the most suitable plant types for specific soil conditions. To establish robust predictive models, we have deployed baseline algorithms such as Random Forest (RF) for building crop yield prediction model and Convolutional Neural Networks (CNN) for soil image recognition model, enhancing our predictive capabilities and enabling more accurate insights.We have developed a sophisticated computer vision-based soil type recognition model, enabling seamless input of soil data through image processing. This user-friendly interface streamlines data acquisition and analysis processes, ensuring accessibility for farmers and policymakers alike to provide them with the optimal crop to be selected to maximize the yield and profit. Through this multifaceted approach, our project aims to empower agricultural stakeholders with invaluable insights, facilitating informed decision-making and fostering sustainable practices in agriculture.

Index Terms– Machine Learning, CNN, Random Forest, Feature Engineering, Computer Vision.

I.INTRODUCTION

In this paper developed a comprehensive web application designed to assist farmers and agricultural stakeholders. Agriculture, one of the most crucial sectors in the world, provides food and a means of subsistence to a few billion people. The agriculture industry does, however, confront various challenges, such as population expansion, resource shortages, and climate change.

The use of machine learning in agriculture has the potential to change crop cultivation and harvesting. For example, predictive models can forecast crop yields, identify disease outbreaks, and recommend the best fertilizers and irrigation practices. These can help farmers save time, reduce costs, and improve their cultivation, thus leading to increased productivity and efficiency in the agriculture industry. In this project, we focus on predicting crop yield using machine learning techniques such as Random Forest, Convolutional Neural Network. By leveraging the power of these models, we aim to develop accurate and efficient tools that can help farmers make informed decisions and improve their yields.

This approach has several key objectives:

- **Enhanced Crop Selection**: By understanding the specific properties of a field through soil image recognition, the system can recommend crops with a higher likelihood of optimal yield in that particular environment.
- **Data-Driven Decisions:** The project seeks to empower farmers with data-driven insights. By combining soil properties with existing crop yield prediction models, farmers can make informed decisions about crop selection, maximizing yield and resource utilization.
- **Improved Efficiency and Sustainability:** Precise crop recommendations based on soil conditions can optimize resource allocation (water, fertilizer) and potentially reduce waste. These advancements contribute to promoting sustainable agricultural practices.
- **Increased Productivity:** Through a combination of accurate yield prediction and tailored crop recommendations, the project aims to contribute to a significant increase in overall agricultural productivity.

II.LITERATURE REVIEW

Anjali Singh et al. [1] discuss the importance of agriculture production and the challenges it faces due to changing climatic conditions. It emphasizes the role of nutrients in crop production and the need for regulated nutrient management to increase crop production.Dr. Reshmi Welekar et al. [2] Examine farming environments and situations to identify the best parameters and data for enhancing crop production, utilizing data mining and machine learning methodologies.Abhinav Sharma et al. [3] explore the use of machine learning in agriculture, focusing on key applications such as supply chain optimization, crop selection, logistics, crop yield prediction, soil data analysis, and weather data analysis.

Feng Liu et al. [4] introduce the application status of machine learning in various fields of agriculture and prospects for the application of machine learning in agricultural production. It also shows how knowledge agriculture improves the sustainable productivity and quality of products to enable plant yield prediction, weed testing, and species testing.Mahendra N et al. [6] explore the application of machine learning in agriculture, focusing on optimizing crop selection and fertilizer use based on soil content and weather parameters to improve crop yield and reduce soil pollution.Ersin Elbasi et al. [5] survey AI applications in the agricultural sector. The discussion begins with foundational knowledge of AI, encompassing all AI techniques employed in the realm of agriculture. This includes machine learning, the Internet of Things (IoT), expert systems, image analysis, and computer vision.

III.PROBLEM STATEMENT

Crop Yield Prediction using Machine Learning is the system that is used to predict the Yield of a crop from the values and images provided by the user. The system processes the input values provided by the user and gives the output as the maximum yielding crop as recommendation, along with the yields of other crops. With an unpredictable condition of nature, accurate analysis benefits the farmers to cultivate the crop based on those conditions. We used CNN and Random Forest algorithms for crop yield predictions and upon comparative analysis Random Forest was giving higher accuracy. The main objective is to develop a reliable system for crop yield prediction with accurate results.

Our project's primary goal is to assist farmers in determining which crop they have to grow to obtain high yield with the predicted conditions in that area. We used CNN (Convolutional Neural Network) which is a Deep Learning algorithm, this is used as it is the best algorithm for image processing for soil-type recognition. We also used Random Forest algorithm for prediction of yield. The steps are providing image to the algorithm to get the type of soil. Based on the location given by the user the model will get that location Humidity, Precipitation, Temperature, Season which are taken as input to the model, based on which the model predicts crop yields of all available crops

IV.METHODOLOGY

Requirements Gathering: Understand the needs of farmers and stakeholders regarding crop yield prediction and agricultural guidance. Identify the features and functionalities required in the application.

Research and Planning: Explore relevant machine learning algorithms for crop prediction, Weather APIs for automation of system, Plan the integration of these technologies into the application.

Data Collection and Preprocessing: Gather agricultural dataset containing features like Area(hectares), Yield(tons),
Temperature (in °C), Precipitation (in mm), Humidity (in g/m^3),
Soil type - The dataset we are using has six different types of soil which are peaty soil, loamy soil, chalky soil, clay soil, sandy soil, silty soil. Soil image dataset obtained from Kaggle[9].
Districts: The dataset also consists of all the districts that are in Maharashtra,
Season: The dataset contains three main crop seasons, namely Rabi, Kharif, and Zaid.
Crop: The dataset also contains crop name features, namely Bajra, Wheat, and Jowar.
Prepare the data by managing null values, transforming categorical variables into numerical form, and normalizing numerical attributes and perform other necessary data preprocessing steps.

Model Development: Develop ML models using Random Forest for crop yield prediction based on preprocessed data and CNN model for soil image recognition and integrate both models. Train the models using historical agricultural data to recognize the soil type and pass the value to the crop yield prediction model to predict the yields of various crops and give the maximum yielding crop as the optimal crop recommendation.

API Integration: Integrate with Open Weather APIs to provide temperature, humidity, precipitation inputs for the crop yield prediction model.

Interface Design: Design a user-friendly interface using the Streamlit framework, allowing farmers to use the automated input option to give only three inputs and manual input option for users with technical data.

Testing and Validation: Test the application for accuracy, reliability, and performance. Validate the predictions and generate responses against known data and expert knowledge. Gather feedback from users and stakeholders for improvements.

Deployment: Deploy the application using Streamlit cloud and publish the web app in a web domain for everyone to access the project from anywhere.
Anyone can Use the deployed app online which is available in the following link:
https://optimal-crop-recommendation-system.streamlit.app/

The working of Random forest algorithms used in this model.

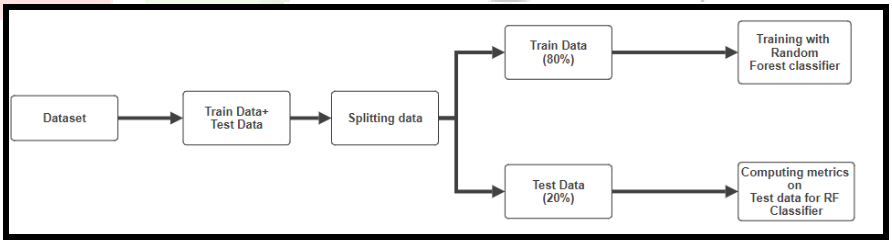

Fig. 4.1: Block Diagram describing the Training of the Machine Learning model.

Working of the crop recommendation system automated in the methodology process:

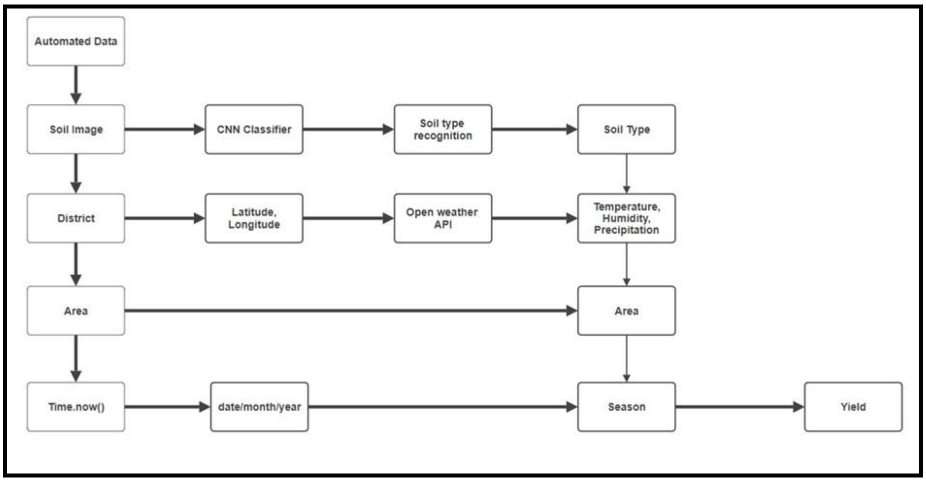

Fig. 4.2: Block Diagram describing the working Automated input Option

Streamlit:

Streamlit[8] is a widely-used Python library that enables the creation of dynamic web applications. In this project, Streamlit is utilized to create a user-friendly interface for crop prediction and translation applications. Streamlit allows developers to write simple Python scripts to create web applications without needing to write HTML, CSS, or JavaScript code.

With Streamlit, developers can easily incorporate data visualization, user inputs, and dynamic outputs, making it ideal for rapid prototyping and deploying machine learning applications.

In this project, Streamlit is used to collect user inputs (such as text queries, audio recordings, or numerical parameters) and display the predicted crop types and translations generated by the model).

Random Forest:

Random Forest[10] is one of the most popular and commonly used algorithms by Data Scientists. Random forest is a widely-used supervised machine learning algorithm that excels in both classification and regression problems. It constructs decision trees on various samples and uses majority voting for classification and averaging for regression. A key feature of the random forest algorithm is its ability to handle datasets with both continuous variables (for regression) and categorical variables (for classification). This versatility allows it to perform effectively in both types of tasks. Random forests are ensemble learning methods that combine the outputs of multiple decision trees to improve predictive performance and reduce the risk of overfitting. By aggregating the predictions of individual trees, random forests enhance accuracy and robustness. They also have built-in mechanisms for estimating feature importance, helping to identify the most significant variables in the dataset. Additionally, random forests are relatively efficient to train and can handle large datasets with high dimensionality.

CNN (Convolutional Neural Network):

Convolutional Neural Networks (CNNs)[7] are a class of deep learning models frequently utilized for tasks such as image recognition, classification, and segmentation.. Inspired by the visual cortex of the human brain, CNNs consist of multiple layers of neurons, including convolutional, pooling, and fully connected layers, designed to extract hierarchical features from input images. CNNs leverage convolutions to detect patterns and spatial dependencies within the data, enabling them to learn representations directly from raw pixel values. Their ability to automatically learn relevant features from data makes CNNs highly effective for tasks like object detection in images, medical image analysis, and autonomous driving. With frameworks like TensorFlow and Keras, building and training CNNs has become more accessible, empowering developers to create sophisticated computer vision applications.

V. RESULTS AND DISCUSSION

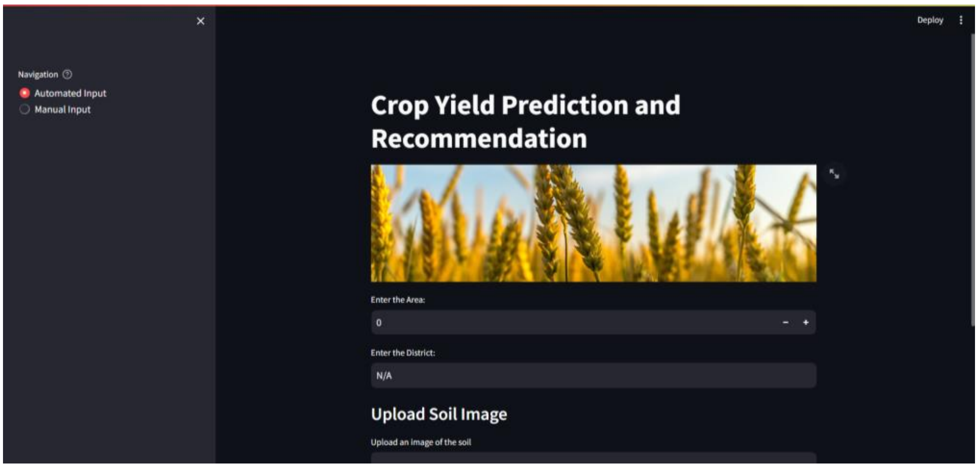

Fig. 5.1: Initial Webpage

The above Fig. 5.1 displays the Home page of the crop yield prediction and recommendation system

AUTOMATED INPUT OPTION

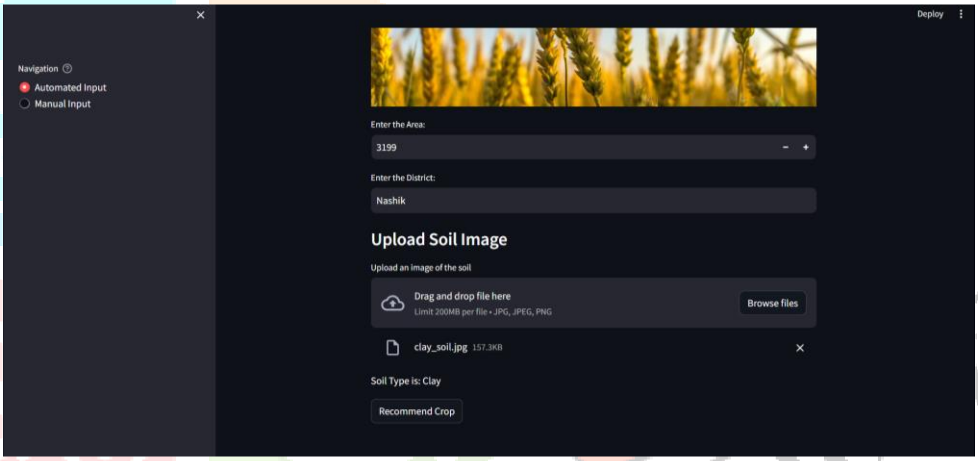

Fig. 5.2: Input

The above Fig. 5.2 displays the Web page for the Automated input option on the sidebar.

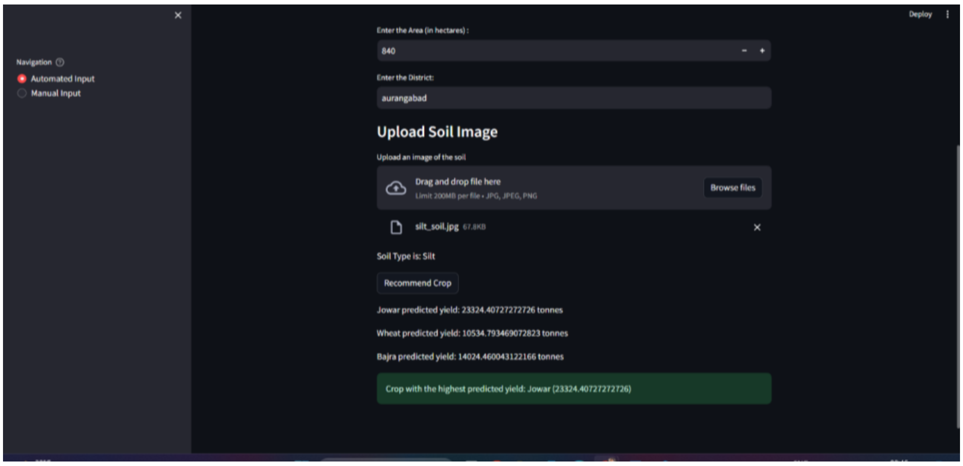

Fig. 5.3: Output

The above Fig. 5.3 displays the Output for the Automated input option after giving input.

MANUAL INPUT OPTION

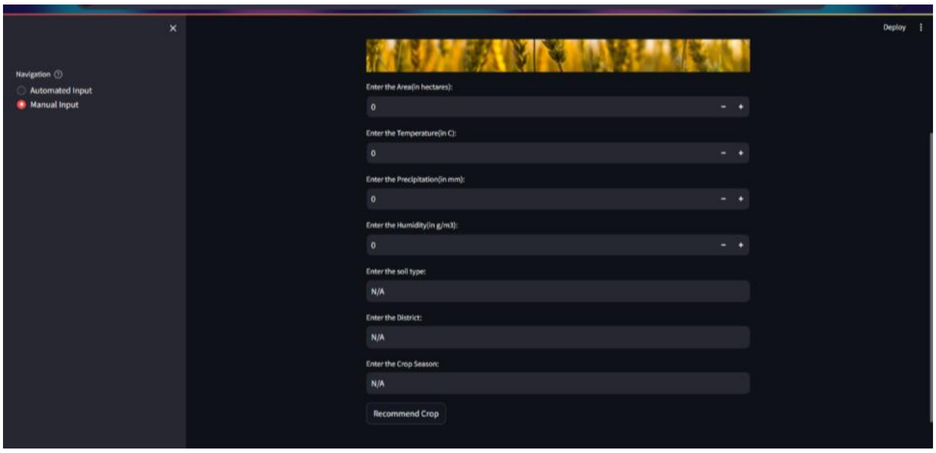

Fig. 5.4: Input

The above Fig. 5.4 displays Web page for the Manual input option on the sidebar

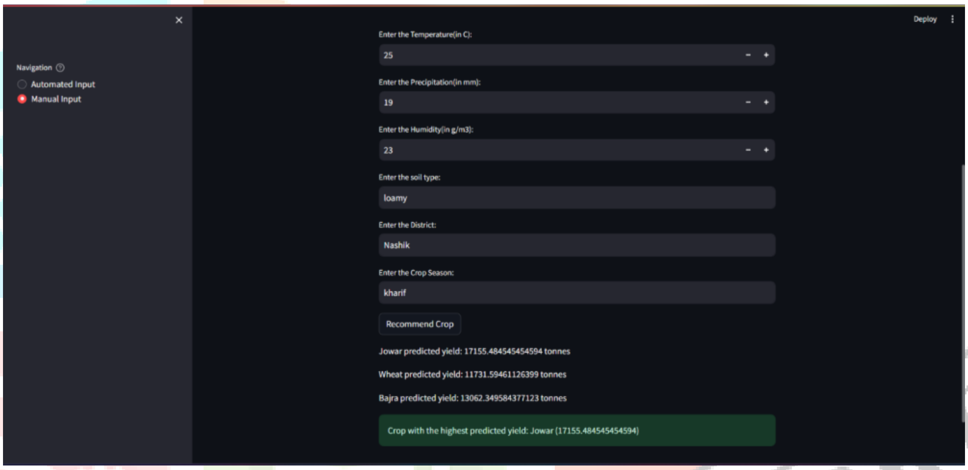

Fig. 5.5: Output

The above Fig. 5.5 displays the Output for the Automated input option after giving input.

VI. CONCLUSION AND FUTURE SCOPE

Conclusion:

In conclusion, our comprehensive agricultural project represents a significant advancement in the realm of crop yield prediction, harnessing the power of advanced machine learning methodologies and soil image recognition. Through the integration of Random Forest (RF) and Convolutional Neural Networks (CNN), we have developed robust predictive models capable of providing valuable insights into optimal plant types for specific soil conditions, thereby empowering farmers and policymakers alike.

It is essential to recognize the inherent limitations of predictive models, as highlighted by the adage, "there are two sides to a coin." While our models offer valuable guidance, it is crucial to understand that no model can be considered infallible. Therefore, farmers should view our predictions as tools to aid in their decision-making processes rather than definitive solutions.

Furthermore, our project emphasizes the importance of transparency and accessibility in agricultural technology. The incorporation of explainable AI techniques ensures that users can understand and trust the insights provided by our models. Additionally, our user-friendly interface, enabled by computer vision-based soil type recognition, facilitates seamless data acquisition and analysis, further enhancing the practicality of our solution.

Future Scope:

Looking ahead, there are several avenues for future development and expansion of our project:

1. Enhanced Model Training: Continuously refining and updating the CNN models with larger and more diverse datasets can improve their accuracy and robustness in soil image analysis.
2. Integration of Remote Sensing Data: Incorporating remote sensing data such as satellite imagery and drone data can provide additional insights into crop health and environmental conditions, further improving yield prediction accuracy.
3. Real-time Monitoring and Decision Support: Developing a real-time monitoring system coupled with decision support tools can enable farmers to receive timely recommendations based on current soil and environmental conditions, facilitating proactive management practices.
4. Expansion to Different Crops and Regions: Extending the application of our model to different crops and geographical regions can cater to the diverse agricultural landscapes and crop varieties across the globe, maximizing its impact and relevance.
5. Collaboration with Agricultural Stakeholders: Collaborating with agricultural researchers, policymakers, and industry stakeholders can facilitate knowledge exchange, validation of models, and adoption of innovative technologies in agriculture

REFERENCES

[1] "Precision Agriculture and Machine Learning" by Anjali Singh, Neha Janu, Sainyali Trivedi, and Mudeet Jain. 978-1-6654-7988-2/22/$31.00, ©2022 IEEE 659| DOI: 10.1109/aic.2022.110

[2]" Optimizing Crop Yield in Agriculture using Data Mining and Machine Learning Techniques" by Dr. Reshmi Welekar, and Dr. Charanjeet Dadiyala(DOI 979-8-3503-3575-0/23/$31.00), ©2023 IEEE

[3] "Development of High-Quality Crops Using an Optimized Machine Learning Agriculture Environment" by G. Venkatakotireddy, C. Lakshminatha Reddy, Ramyadevi. R., J. Prabhakaran, T. Nivethitha, Mahaveerakannan R., DOI: 10.1109/ICAIS56108.2023.10073727, ©2023 IEEE

[4] "The Application of Machine Learning and IoT in Smart Agriculture" by Feng Liu andXiaolu Yan.DOI: 10.1109/GCRAIT55928.2022.00019, ©2022 IEEE

[5] "Artificial Intelligence Technology in the Agricultural Sector: A Systematic Literature Review" by Ersin Elbasi, Nour Mostafa, Zakwan AlArnaout, Aymen I. Zreikat, Elda Cina, Greeshma Varghese, Ahmed Shdefat and Ahmet E. Topcu.DOI: 10.1109/ACCESS.2022.3232485, ©2022 IEEE

[6] "Crop Prediction using Machine Learning Approaches" by Mahendra N, Dhanush Vishwakarma, Nischitha K, Ashwini, andManjuraju M. R., DOI: 10.17577/IJERTV9IS080029, ©2020 IJERT

[7] "https://www.tensorflow.org/tutorials/images/cnn"

[8] "https://docs.streamlit.io/"

[9] "https://www.kaggle.com/datasets/jhislainematchouath/soil-types-dataset"

[**10**] "https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html**"**