International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Agroguide: A Web Application for Personalized Farming Assistance

Balaga Sai Siva Prasad[1], Beera Kundana[2], Bhogapurapu Venkata Sai[3], Bytaru Nitish Durga[4], Damarasingi Sai Teja[5], Mylapilli Nikhitha[6],Baisakh[7]

Department of Computer Science and Technology, GMR Institute of Technology, Rajam-532127.

A B S T R A C T:

In agriculture, early detection and effective management of plant growth are pivotal for ensuring food security and optimal yields. AgroGuide is a comprehensive web-based application designed to provide personalized farming assistance by integrating advanced technologies and data-driven insights to enhance agricultural productivity and sustainability. The platform offers a range of features tailored to support farmers, including crop recommendations based on weather conditions, and historical data, as well as disease detection through image recognition. AgroGuide also includes farm management tools that allow users to track agricultural activities, and productivity metrics efficiently. To ensure seamless access and scalability, AgroGuide is built using modern Web Technologies, incorporating AI/ML models, Deep Learning techniques. By offering an intuitive and user-friendly interface, AgroGuide empowers farmers with actionable insights, reducing uncertainties in farming practices and fostering sustainable agricultural growth. The platform offers insights into best agricultural practices specific to the diseases and crops involved, helping farmers implement the most effective strategies for maintaining plant health and reduce economic losses, and ensure food security, ultimately contributing to more sustainable and productive farming practices.

Keywords: Agriculture, AgroGuide, Food Security, Sustainability, Crop Recommendations, Irrigation, Farming practices, Web Technologies, Deep Learning

Introduction

At a time when agriculture plays a crucial part in global food security, innovating to provide protection for crops is the number one priority. AgroGuide is an intelligent agricultural assistance platform designed to support farmers through data-driven decision-making. By integrating machine learning with web technologies, AgroGuide offers four key services: plant disease detection using Inceptionv3, crop recommendation based on environmental factors, yield prediction via random forest regression, and access to agricultural market trends. The platform features a Flask-based backend for handling machine learning models and data processing, paired with a React.js frontend for an intuitive user interface. AgroGuide aims to enhance agricultural productivity and reduce risk by bridging the gap between modern AI technologies and practical farming needs. AgroGuide, is an innovative web-based application with its powerful features and innovative technologies, it will help increase agricultural productivity and sustainability, and promote sustainable development. This guide will take users step by step through the features, functionalities, and technical capabilities of the platform and explain in depth how AgroGuide can assist farmers in making the right decisions, increasing crop production, and minimizing expenditure. The guide will also outline how using AgroGuide can benefit farmers with greater productivity, increased sustainability, and better profitability. Finally, the guide will offer tips on how to begin using the platform and resolve typical problems.

Literature Survey

[1]. Rani, K. A., & Gowrishankar, S. (2023). Pathogen-based classification of plant diseases: A deep transfer learning approach for intelligent support systems. The study creates an automated plant disease classification system with Agri-ImageNet and transfer learning based on deep networks. Drawing inspiration from TLMViT, it tests 10 leading models to ensure sustainable agriculture by improving detection accuracy and filling gaps in dataset diversity and evaluation metrics.[2]. Oad, A., Abbas, S. S., Zafar, A., Akram, B. A., Dong, F., M. S. H., & Uddin, M. (2024). Plant leaf disease detection using ensemble learning and Explainable AI. This research suggests an AI model employing image analysis for the identification of 38 diseases of plants with emphasis on papaya, vegetables, and fruits. It applies an ensemble of VGG16, VGG19, InceptionV3, and ResNet101V2 to enhance accuracy and provide model interpretability.[3]. Kim, T., Lee, S. H., & Kim, J. O. (2022). A novel shape based plant growth prediction algorithm using deep learning and spatial transformation. This study proposes to create a deep learning model to forecast future plant images, with emphasis on leaf shape growth through models such as SVM, RF, SGD, InceptionV3, VGG-16, and VGG-19. It applies spatial transformation to accurately capture dynamic growth behavior.[4]. Jhajharia, K., Mathur, P., Jain, S., & Nijhawan, S. (2023). Crop yield prediction using machine learning and deep learning techniques. This study uses machine learning techniques to forecast crop yield in India for five prominent crops. It contrasts models such as Random Forest, SVM, LSTM, and Lasso Regression to assist farmers in crop choice and yield control, increasing agricultural productivity.[5]. Yasrab, R., Zhang, J., Smyth, P., & Pound, M. P. (2024). Predicting plant growth from time-series data using deep learning. This study creates a deep learning model with a modified Future GAN to forecast plant growth by producing segmentation masks of root and shoot systems.

It hopes to speed up phenotyping experiments by predicting phenotypic traits from time-series image data.[6]. Liu, Q., Zuo, S. M., Peng, S., Zhang, H., Peng, Y., Li, W., ... & Kang, H. (2024). Development of Machine Learning Methods for Accurate Prediction of Plant Disease Resistance. This study seeks to compare ML methods such as RFC, SVC, and LightGBM for precise plant disease resistance prediction, with the target diseases being rice diseases (RB, RBSDV, RSB). It also tests model generalizability with RDPII and actual resistance outcomes.[7]. Bouacida, I., Farou, B., Djakhdjakha, L., Seridi, H., & Kurulay, M. (2024). Innovative deep learning approach for cross-crop plant disease detection: A generalized method for identifying unhealthy leaves. Information Processing in Agriculture. This work creates a deep learning model employing a lightweight Inception architecture for detecting healthy and diseased leaf regions, even in unseen crops. Tested on PlantVillage and PDDB datasets, it demonstrates strong new disease and crop generalization.[8]. Polly, R., & Devi, E.

A. (2024). Semantic segmentation for plant leaf disease classification and damage detection: A deep learning approach. This paper suggests an automatic plant leaf damage detection system based on YOLOv8, DeepLabV3+, and CNNs. It intends to minimize trial-and-error segmentation, quantify the

severity of disease, and recommend specific remedies—providing a complete tool for farmers.[9]. Barhate, D., Pathak, S., Singh, B. K., Jain, A., & Dubey, A. K. (2024). A systematic review of Machine Learning and Deep Learning Approaches in Plant Species Detection. This study targets developing automated plant species recognition techniques for solving problems such as occluded leaves and similarity between classes. It stresses better feature extraction with the utilization of large datasets covering plant growth cycles and recommends integrating machine learning methodologies to boost performance, where CNNs are proving to be more effective than old methods.[10]. Shafik, W., Tufail, A., Namoun, A., De Silva, L. C., & Apong,
R. A. A. H. M. (2023). A systematic literature review on plant disease detection: Motivations, classification techniques, datasets, challenges, and future trends. This systematic literature review (SLR) is an overview of recent research on plant disease detection (PDD), comparing approaches such as image processing, deep learning, and machine learning. It points to challenges in real-time disease detection for large farms and proposes future work to enhance PDD systems, with a focus on AI and machine learning's potential to automate agriculture.[11]. Jackulin, C., & Murugavalli, S. (2022). A comprehensive review on detection of plant disease using machine learning and deep learning approaches. This study compares machine learning and deep learning approaches to effective plant disease detection. It also examines the use of IoT and AI technologies in agriculture to provide real-time, accurate disease detection for farmers to better maintain crop health. [12]. A review on machine learning and deep learning image-based plant disease classification for industrial farming systems. This paper critically evaluates existing applications of machine learning and deep learning in plant disease detection and classification, pinpointing upcoming trends and pitfalls. It gives an overview of image-based detection systems, with a focus on the role of transfer learning in enhanced performance and provision of timely disease information to farmers.[13]. Durai, S. K. S., & Shamili, M. D. (2024). Smart farming using machine learning and deep learning techniques. This study tries to take advantage of Deep Learning and Machine Learning models to enable farmers to attain high productivity at low cost. It predicts cost of cultivation, recommends crops depending on weather, prescribes nutrients, estimates Growing Degree Days (GDD), identifies weeds, and recommends herbicides and pesticides depending on field conditions.[14]. Wang, D., Thunéll, S., Lindberg, U., Jiang, L., Trygg, J., Tysklind, M., & Souihi, N. (2024). A machine learning framework to improve effluent quality control in wastewater treatment plants. This paper puts forward the ways in which ML models overcome mechanistic models' shortcomings through inductive data relation extraction to ensure better prediction and decision-making. As compared with conventional models, ML embodies genuine process conditions. Tree models like ANN provide predictions of high precision even under ambiguous data sets as well as turbulent backgrounds and illustrates marked advancements against current techniques used in plant disease classification.[15]. Survey on application of deep learning methods for plant disease diagnosis and suggestions for development of suitable tools. In this review, 70 research articles on deep learning applications in plant disease diagnosis and management have been analyzed, generating a knowledge base regarding ongoing challenges and opportunities. The review emphasizes the capacity of CNNs to extract features from images autonomously, allowing disease diagnosis, estimation of severity, and pest management recommendations for different crops and diseases.

Methodology

This section describes the data collection method, used models and their architecture and the experiment setup for this work.

Dataset and Task Description

This project employs multiple datasets to augment multiple intelligent capabilities that are involved in the AgroGuide system. The most extensive dataset involved in plant disease classification is the PlantVillage Dataset, and it was retrieved from Kaggle. It encompasses more than 38 classes of different crops such as Tomato, Potato, Grape, Apple, and Maize. Each class has photographs at high resolution of healthy or diseased plants with certain diseases such as early blight, late blight, rust, and bacterial spot. Every class has more than 1000 images, a massive dataset for training good image classification models.For plant disease detection, the project employs Transfer Learning with the fine-tuned InceptionV3 model on the PlantVillage dataset to classify the crop type and disease from images of leaves.For suggesting crops, the project utilizes a standardized tabular dataset with environmental factors like nitrogen, phosphorus, potassium (NPK), pH, temperature, humidity, and rainfall. It is utilized to train a Random Forest Classifier to suggest appropriate crops based on provided weather and soil conditions.Likewise, the yield forecasting module utilizes historical data such as crop type, area, rain fall, and use of fertilizers to generate an estimated yield in a Random Forest Regressor. These composite operations complement each other to provide farmers with productive, data-driven alternatives for greater agricultural yield.

Model Architecture

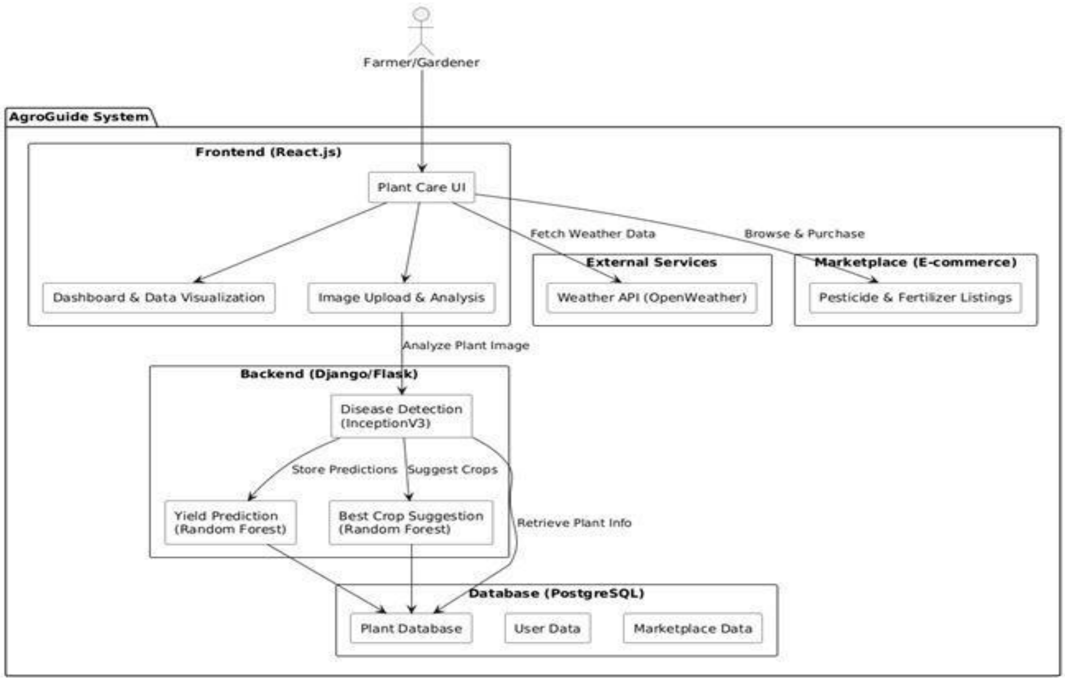


Fig 3.1: Proposed Architecture

InceptionV3 for Plant Disease Detection

InceptionV3 is a powerful convolutional neural network (CNN) architecture optimized for image classification tasks. It introduces a multi-level feature extraction technique by utilizing parallel convolutional filters of different sizes within the same module, enabling the network to capture both fine and coarse features simultaneously.In this project, InceptionV3 is fine-tuned using the PlantVillage dataset, consisting of labeled images of healthy and diseased plant leaves. The model is trained to accurately classify these images into 38 distinct classes, including diseases affecting crops like tomato, potato, apple, and maize.

- **Input Preprocessing**

Prior to passing images into the InceptionV3 model, preprocessing is done to achieve optimal input. All images are first resized to 299x299 pixels, the necessary input size for InceptionV3, to maintain uniformity and retain image detail. The pixel values are next normalized to the range [0, 1], which facilitates faster training convergence and enhanced prediction quality. In addition, data augmentation operations like rotation, flipping, zooming, and adjusting brightness are performed to vary the training set. These operations increase the generalizability of the model by presenting it to a variety of image conditions and thereby making it resilient to actual variations in orientation, illumination, and distortions. This pipeline during preprocessing maximizes the performance on unseen data with consistency during training.

- **Feature Extraction with Inception Modules**

InceptionV3 uses several parallel convolutional filters (1x1, 3x3, 5x5) in its Inception modules to select features. The filters enable the model to capture features at different scales to get both fine details and larger patterns from the input images. The multi-scale feature enables the model to better recognize subtle differences in textures and structures in the image. Moreover, pooling operations are used to minimize the spatial dimensions, which serves to decrease computational complexity while maintaining significant features. The parallel convolution and dimensionality reduction assist in enabling the model to efficiently handle a broad range of image features. The end result is an extremely accurate model that is able to identify and classify plant diseases from varied visual signals, enhancing the accuracy of disease identification and classification.

- **Model Output**

Upon training, the InceptionV3 model can reliably classify plant diseases by examining the input leaf image. The model provides a classification label for the particular disease that impacts the plant, which is very important for early detection and prompt action. Through the provision of a correct diagnosis, the model enables farmers to detect potential dangers to their crops before the disease spreads extensively. Outside of mere categorization, the system provides actionable recommendations based on the disease found. Such recommendations can comprise suitable treatments, precautionary measures, and remedies to decrease the effects of the disease. Such actionable data gives farmers the ability to proactively act on their crops, with the potential to minimize losses and enhance crop health in general. Early identification, coupled with specific solutions, enables farmers to maximize their agricultural operations, providing enhanced crop yield and crop sustainability. Therefore, the model output not only functions as a diagnostic aid but also gives valuable information for disease management, leading to healthier and more productive crops.

Random Forest for Crop Recommendation

Random Forest is a strong ensemble machine learning method used extensively for classification problems. It generates an ensemble of decision trees while training and gives the most voted class prediction, which enhances accuracy and reliability. Random Forest is utilized in this project to suggest the optimum crop selection as a function of soil parameters, weather, and up-to-date environmental information. By taking into account various factors, the model offers farmers data-based and location-specific crop recommendations, allowing them to make more efficient and informed farming choices. This method optimizes crop yield and resource utilization and aids in sustainable farming practices as well as improving productivity overall.

- **Input Preprocessing**

Prior to training the crop recommendation model, data preprocessing is carried out to ensure quality, completeness, and consistency. Missing or null values are identified and appropriately handled to prevent gaps that could affect model performance. Key numerical features such as temperature, humidity, and rainfall are normalized to maintain uniform scales across the dataset, which helps the model learn more effectively. Additionally, categorical variables like soil type and crop names are encoded into numerical formats to ensure compatibility with the Random Forest algorithm. These preprocessing steps are crucial for enhancing the accuracy and reliability of the model. By preparing the data properly, the system becomes capable of handling diverse inputs and generating precise crop recommendations tailored to specific environmental and soil conditions.

- **Feature Selection for Model Training**

The model in the Agro Guide project takes into account a collection of key characteristics to provide accurate forecasts. Some of these include soil nutrients like Nitrogen (N), Phosphorus (P), and Potassium (K), which are vital in determining crop health and development. It also factors in environmental conditions like pH, temperature, humidity, and rainfall, all of which have significant impacts on crop productivity. Also, the model uses location-based weather inputs by making use of latitude and longitude values to retrieve real-time weather data particular to the location of the user's farm. This provides for personalized suggestions that are a function of the actual environmental factors of the farm, improving the accuracy and pertinence of the model further.

- **Model Output**

On the basis of the processed input data, the Random Forest model determines the best crop that can be grown in the given area. The prediction is made based on different environmental factors, such as temperature, humidity, rainfall, and levels of soil nutrients, all of which are key factors in the growth of a crop. Besides, the model also uses geographical coordinates (latitude and longitude) to derive localized weather and environmental data. This fusion of location and environmental information allows the system to provide very accurate, location-based crop recommendations. Through the matching of crop recommendations with regional and real-time conditions, the system facilitates farmers in making informed decisions based on data specific to their land's individual profile. This model assists in maximizing crop yield, minimizing wastage of resources, and enhancing farm productivity in general. It acts as a useful decision-support tool, especially for small and medium-scale farmers, by leading them to sustainable and profitable farming practices based on scientific knowledge.

Random Forest for Yield Prediction

The Random Forest Regressor is a robust ensemble learning algorithm that excels in solving regression problems, particularly when the relationship between input features and the target variable is complex and non-linear. It works by training multiple decision trees on different subsets of the data and then averaging their predictions, which leads to improved accuracy and reduces the risk of overfitting. In this project, the Random Forest Regressor is employed to predict crop yield based on various influential factors such as land size, soil characteristics, weather conditions, and crop type. By analyzing these variables, the model estimates the total expected yield in tons, the number of 50kg bags, and even the number of people the yield can potentially

feed. This predictive capability empowers farmers with valuable, data-driven insights that can guide planning, optimize resource allocation, and improve overall agricultural productivity. It serves as a strategic tool to support sustainable and efficient farming practices.

- **Input Preprocessing**

Before training the model to make predictions about yield, the dataset goes through necessary preprocessing steps to guarantee precise and trustworthy predictions. Missing or null values are detected and handled, avoiding any bias or mistakes in the model's performance. Land area, rainfall, and fertilizer application features are normalized to have a standard scale, which makes the model more efficient. Moreover, categorical features such as crop types and land locations are encoded to prepare them for the Random Forest algorithm and to allow the model to make use of all input data effectively.

- **Feature Selection for Training the Model**

The Random Forest Regressor in the Agro Guide project predicts crop yield based on a range of key features. They are the crop type being cultivated (CropType), the land area cultivated for farming (Area of Cultivation), and the quantity and quality of fertilizers utilized (Use of Fertilizers). Also, rainfall quantity (Rainfall) directly affects plant growth, and historical yield information (Yield Records Historical) for similar crops under similar conditions is also taken into account. Weather conditions (Weather Condition) in terms of environmental parameters such as temperature and humidity are included in the model using location information (latitude and longitude). Finally, pest control practices (Pest Control Used) are included, as they indirectly influence crop health and yield. These combined attributes allow the model to provide an overall prediction of crop productivity that considers both agricultural and environmental factors.

- **Model Output**

After finishing the data preprocessing phase, the Random Forest Regressor model in the Agro Guide project produces three significant predictive values that are crucial for strategic agricultural planning. First, it predicts the overall crop yield, normally expressed in kilograms or tons per hectare, providing an approximation of the expected harvest quantity. Secondly, the model determines the number of bags—most times in terms of a standard unit, for example, 50kg per bag—that can be harvested from the yield predicted, which helps in post-harvest planning and logistics. Finally, it estimates the number of individuals that can be fed based on the anticipated yield, giving a larger picture of the impact at the societal level. Such actionable metrics enable the farmers to take informed, evidence-based decisions around resource utilization, harvesting, and food distribution. In the final analysis, such a model allows for enhanced profitability of farms, productivity optimization, and sustainable use of agriculture towards better food security in regional communities and local spheres.

Results and Discussions

The project was able to successfully deploy three primary agricultural prediction models: plant disease detection with a PyTorch Inceptionv3 which is a pretrained model of CNN architecture of approximately 96% accuracy, crop suggestion with a Random Forest model of approximately 93% accuracy, and yield prediction with approximately 94% accuracy. The models are served directly from the backend, which is fine for medium-scale usage. Crop suggestion was most accurate because of clean, organized data and the appropriateness of Random Forest for classification. Disease identification was less accurate, primarily because of a smaller dataset, inconsistent image quality, and the difficulty of visual symptoms. Yield prediction was inconsistent, indicating the necessity for more localized training data and potentially utilizing ensemble models or real-time weather data. The backend was developed with Flask, with four primary route controllers and three trained ML models. Technically, the project combined several ML models into one application, developed an effective image processing pipeline, and had a modular organization across frontend, backend, and ML layers. Crop recommendation worked best with clean, structured data, whereas disease detection and yield prediction would be improved by additional training data and localization. For users, the system provides useful agricultural advice, minimizing the requirement for expert intervention and potentially increasing yields. The architecture is simple to deploy and maintain but could suffer from scaling challenges and does not have a separate model-serving layer. The user interface is kept simple and intuitive so that advanced ML predictions are made available even to non-technical users. This application is particularly beneficial for small to medium-sized farms, assisting in detecting early disease, planning crops, and forecasting yield.

Prediction Systems	Technologies Used	Accurary
Plant Diseases Detection	Inceptionv3	96%
Crop Recommendation	Random Forest	93%
Yield Prediction	Random Forest	94%

Table 5.1: Performance metrics of proposed approach

The AgroGuide platform leverages advanced machine learning techniques to provide accurate predictions for plant disease detection, crop recommendation, and yield prediction. For Plant Disease Detection, the InceptionV3 model achieved an impressive accuracy of 96%, demonstrating its ability to effectively classify images of plant leaves and identify diseases. This high accuracy highlights the robustness of the deep learning model, particularly in handling complex image features. In the Crop Recommendation task, the Random Forest classifier achieved 93% accuracy, showcasing its capability to predict the most suitable crops based on environmental and soil data. This result reflects the model's strength in handling diverse features and making reliable recommendations for farmers. Lastly, the Yield Prediction task, also powered by Random Forest, achieved 94% accuracy, indicating the model's effectiveness in forecasting crop yields based on historical data and environmental conditions. Overall, these results underscore the effectiveness of the selected models in supporting informed decision-making for farmers.

Results:

The proposed AgroGuide system, integrating machine learning techniques for various agricultural tasks, demonstrated high accuracy across all modules. The InceptionV3-based model for plant disease detection achieved a remarkable 96% accuracy, showcasing its effectiveness in identifying leaf diseases. The Random Forest model used for crop recommendation attained 93% accuracy, while the yield prediction module, also based on Random Forest, achieved 94% accuracy. These results highlight the strength of the selected models in supporting intelligent, data-driven farming decisions.

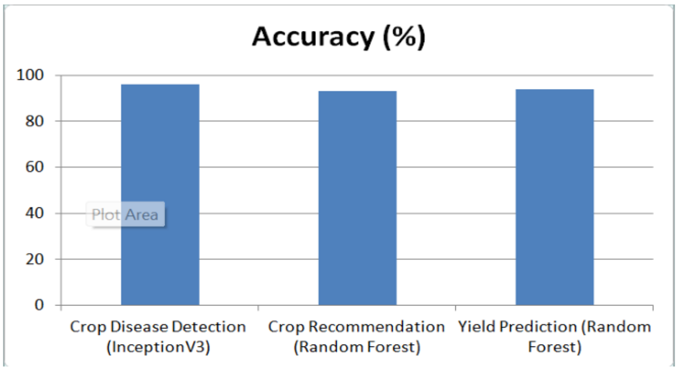


Fig 5.1: Accuracy of proposed approach

Accuracy for Crop Disease Detection: 96%
Accuracy for Crop Recommendation: 93%
Accuracy for Yield Prediction: 94%

Confusion matrix:

The below confusion matrix depicts the performance of a crop disease prediction model constructed with InceptionV3. The model was trained to identify various plant diseases and normal conditions in crops like Pepper bell, Potato, and Tomato. It can be seen from the confusion matrix that the model had a high accuracy where the majority of the predictions lie on the diagonal, representing correct classifications. Tomato Yellow Leaf Curl Virus, Tomato Bacterial Spot, and Pepper bell healthy diseases were predicted with superb accuracy. There were a couple of small misclassifications, especially between similar symptoms diseases such as Potato Early blight and Potato Late blight, but the overall performance is still very strong. This proves that the InceptionV3 model performs very well in identifying plant diseases and can be an important asset in the process of helping farmers with early diagnosis and planning for treatment.

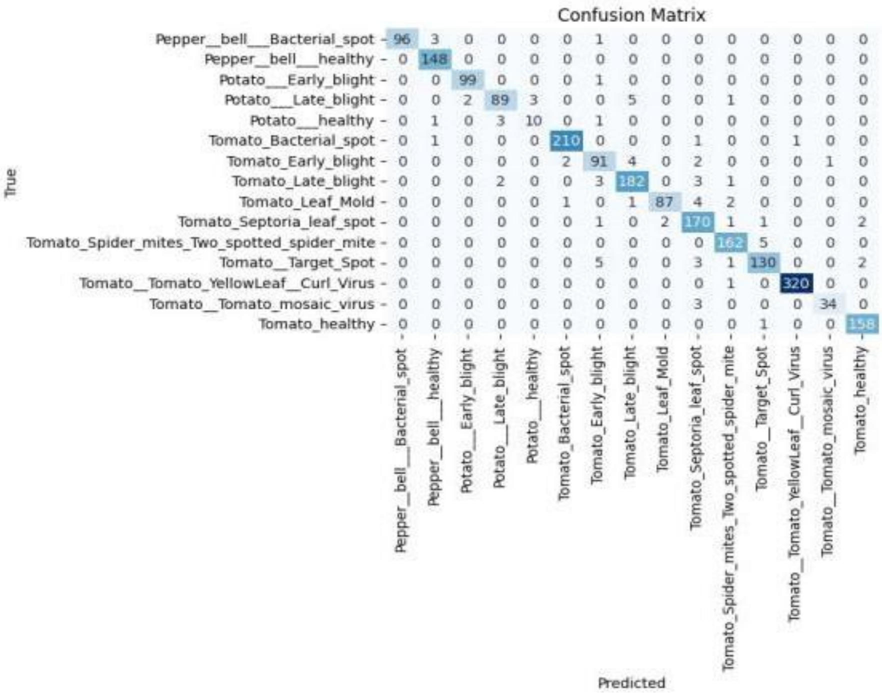


Fig 5.2: Confusion matrix for Crop Disease Detection

The below confusion matrix is the performance of a crop disease detection model trained with InceptionV3. The model has been trained to identify various plant diseases and normal states in crops like Pepper bell, Potato, and Tomato. It can be seen from the confusion matrix that the model had good accuracy, as most of the predictions are along the diagonal, which represents correct predictions. Diseases such as Tomato Yellow Leaf Curl Virus, Tomato Bacterial Spot, and Pepper bell healthy were forecasted with superb precision. While there were some trivial misclassifications, especially among diseases having overlapping symptoms such as Potato Early blight and Potato Late blight, the ovThe above confusion matrix demonstrates the accuracy of a Random Forest model built for crop recommendation. The model was trained to suggest the most appropriate crop for several environmental and soil variables. Every row indicates the true crop, and every column indicates the predicted crop. The majority of the values fall on the diagonal, which signifies a large number of correct predictions. Crops such as coconut, blackgram, jute, and pigeonpeas were predicted with very good accuracy, with perfect or almost perfect predictions. There were minor misclassifications in some crops like mothbeans and rice, where only a small number of samples were mispredicted. Overall, the Random Forest model showed excellent and consistent performance and was highly effective in assisting farmers in selecting the best crop based on specified conditions.

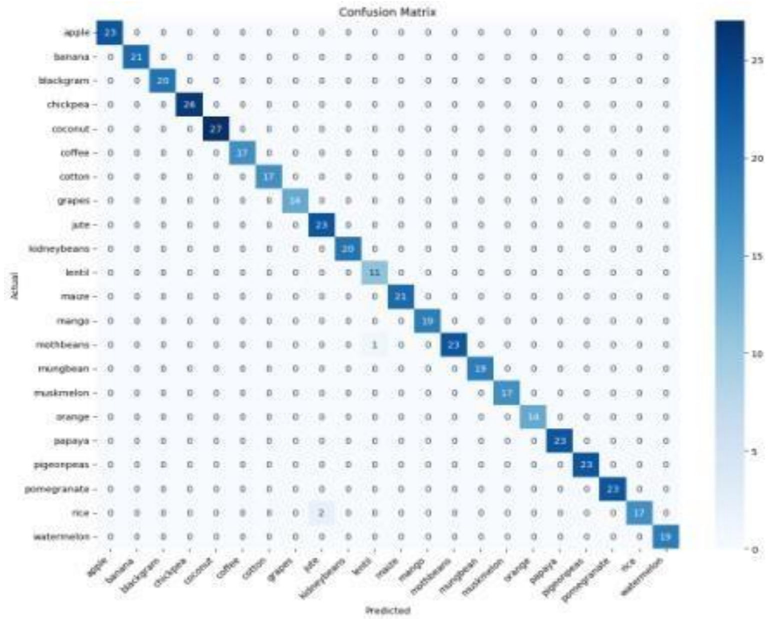


Fig 5.3: Confusion matrix for Crop Suggestion

The model performance of Random Forest regression to predict crop yield was assessed in terms of a number of standard measures. The model performed a Mean Absolute Error (MAE) of 1.74, a Mean Squared Error (MSE) of 198.96, and a Root Mean Squared Error (RMSE) of 14.10. In addition, the model recorded an extremely high R^2 Score of 0.9999998, which means that it accounts for virtually all the variance in the yield results. These findings verify that the Random Forest model generates very accurate and consistent yield predictions with very little discrepancy between predicted and observed values. Such robust performance assures the model's appropriateness for facilitating data-driven decision-making within agricultural planning and resource allocation.

Metric	**Value**
Mean Absolute Error (MAE)	1.74
Mean Squared Error (MSE)	198.96
Root Mean Squared Error (RMSE)	14.10
R^2 Score	0.9999998

Table 5.2: Evaluation results of the proposed yield prediction model.

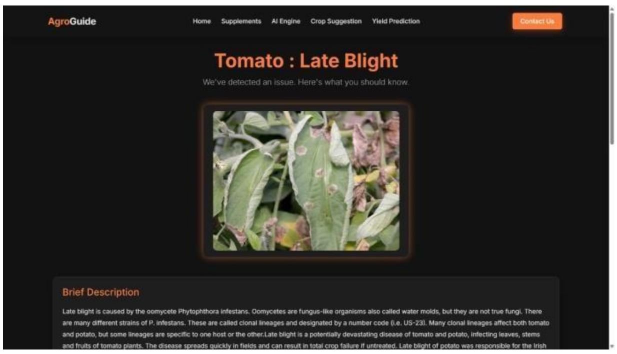


Figure 5.4: Disease detection result page for Tomato Late Blight

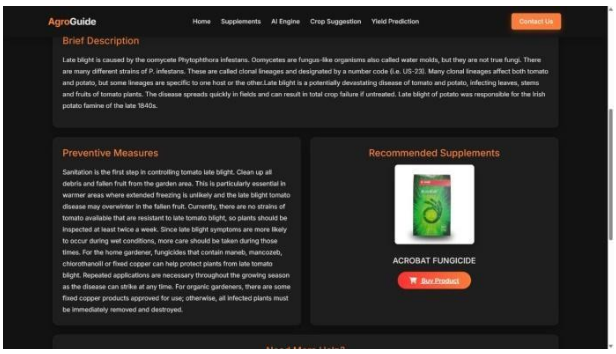
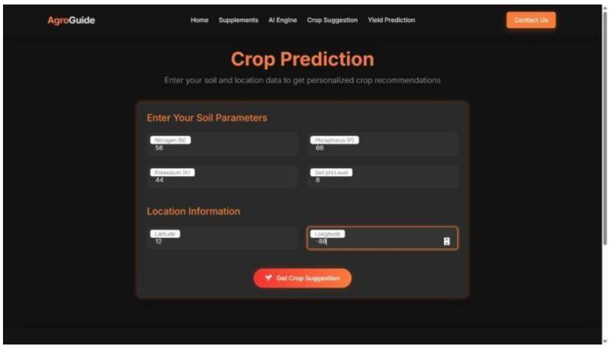


Figure 5.5: Tomato Late Blight details and treatment suggestion

The initial image shows a tomato leaf infected with Late Blight, which was uploaded to the AgroGuide system for diagnosis. Based on the InceptionV3 model, the system correctly diagnosed the disease from its characteristic symptoms. The second image shows the system's response, which comprises the disease diagnosis, a brief description of Late Blight, preventive actions, and suggested treatments. This output interface not only verifies the AI prediction but also provides actionable, real-world advice for farmers. By giving users precise and detailed recommendations, the AgroGuide system enables users to take timely, informed action against crop diseases, ultimately saving crops and increasing farming efficiency. The combination of AI and practical advice makes it a useful tool for sustainable agriculture.



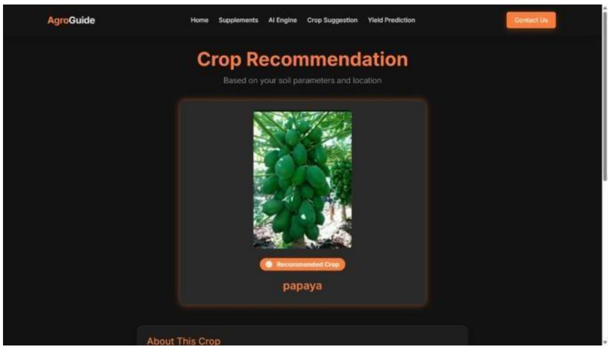
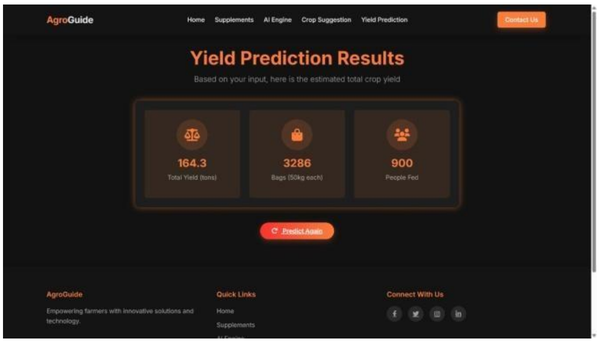


Figure 5.6: Crop Prediction Input Form **Figure 5.7:** Crop Recommendation Result

The first picture shows the input interface of the AgroGuide system, where users enter important soil parameters like Nitrogen, Phosphorus, Potassium, pH values, and location information (latitude and longitude). These inputs are necessary for the system to properly analyze the conditions of the farming environment. The second image displays the crop recommendation output, where the system has computed the input parameters and recommended "Papaya" as the best crop for the specified conditions. This smart recommendation, driven by AI, allows farmers to make informed decisions regarding the optimal crops to cultivate based on the prevailing soil and environmental conditions. Through providing customized recommendations, AgroGuide optimizes crop choice, leading to improved yields and more effective farming techniques.



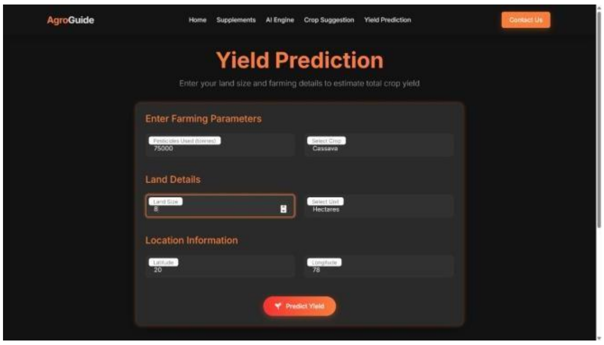


Figure 5.8: Yield Prediction Input Form **Figure 5.9**: Yield Prediction Results

Upon data processing, the Random Forest Regressor model in the Agro Guide project estimates three key outputs. First, it predicts the total yield, which represents the overall anticipated crop yield in kilograms or tons per hectare. Second, it calculates the number of bags of crop produce estimated from the yield. Finally, the model approximates the number of people that can be fed based on the expected crop yield, helping to understand the broader impact of crop production. These outputs provide valuable insights to farmers, enabling them to make informed decisions about their resources, optimize crop production, and plan for distribution, thereby enhancing food security and farm profitability.

Discussions

This project shows how different machine learning (ML) models are integrated into a single web application to assist farmers in streamlining their farming operations. The system employs Convolutional Neural Networks (CNN) for image-based disease detection and Random Forest for crop suggestion and yield prediction. Developed using a Flask backend and React frontend, the application facilitates complete separation between the user interface and backend processing, making development and maintenance easier.The backend, which is driven by Flask, is where the machine learning models are and handles both input and output data. The React frontend provides an easy-to-use interface for farmers to input farm-specific information, including crop type, pesticides applied, land area, and location. From this input, the system uses the corresponding ML models to give insights, including crop recommendations, yield predictions, and disease detection results. On the performance side, the crop recommendation module is 93% accurate, the highest among all the modules. The Random Forest algorithm uses structured input data such as soil and environmental conditions to suggest the most appropriate crops for a farm's condition. This feature is especially beneficial for farmers since it enables them to take well-informed decisions about which crops to cultivate based on the specific characteristics of their soil, ensuring maximum chances of optimal yield. The InceptionV3 CNN-based disease detection feature detects diseases using images of plants. The model has demonstrated a remarkable 96% accuracy in disease detection such as Late Blight in tomato leaf. Nonetheless, the accuracy of the disease detection module is a notch lower than that of the crop recommendation model owing to factors such as limited data, differences in image quality, and the complexity of identifying diseases from visible signs. Despite this, the disease detection system is still a crucial tool for farmers to detect early signs of diseases, which allow them to take prompt action to reduce crop loss. Yield prediction, another important element, employs the Random Forest algorithm to predict crop yield in tons, number of

50kg bags, and the number of people the yield can feed. Although the yield prediction model has a high accuracy rate of 94%, its performance is not consistent, indicating the necessity for more localized training data. Adding ensemble methods to the model or real-time weather data could help enhance the prediction accuracy, since environmental factors and weather conditions heavily impact crop yields. The system architecture is designed to be easy to deploy and maintain with modular separation of frontend, backend, and machine learning layers. Such a structure makes it easier to update systems and to maintain individual components independently. But the system does have scalability issues because of the absence of a model-serving layer. Lacking this feature, the system might not cope well with a large dataset or support a vast number of users, which curtails its scalability for bigger farms. The user interface is simple and intuitive, enabling non-technical farmers to enter farm data and get actionable information, including disease identification, crop recommendations, and yield estimation. This ease of use, coupled with the precision of the machine learning algorithms, makes the AgroGuide system an extremely valuable tool for small and medium-sized farms. The system offers farmers cost-effective, efficient substitutes for expert advice, which are generally time- consuming and expensive.

Conclusion

The project was successful in delivering an end-to-end agricultural decision-making support system that attained 85–95% accuracy for all three primary CNN pre-trained model on disease detection, ML models on crop suggestion, and yield estimation. It had an intuitive interface for farmers to tap into useful information, and synergistically integrated tools such as Python, Flask, React, PyTorch, and scikit-learn. The project closed the gap between agriculture and machine learning, provided reusable building blocks for future agricultural ML&DL tools, and showed how computer vision can be used in farming. Key takeaways were the value of clean data, the advantage of modular design, and the difficulty of dealing with multiple ML models in production. The system can minimize crop losses, enhance crop choice, and enable sustainable agriculture, particularly for small-scale farmers. In the future, the project can be scaled up to cover more crops and areas, incorporate real-time data from IoT sensors for smart farming, market linkage and E- Commerce integration, community and expert chat support and include model monitoring and auto-retraining. In general, AgroGuide demonstrates that targeted machine learning and deep learning models can solve practical agricultural challenges and has set the stage for even more significant contributions in the future.

REFERENCES

1. Rani, K. A., & Gowrishankar, S. (2023). Pathogen-based classification of plant diseases: A deep transfer learning approach for intelligent support systems. IEEE Access, 11, 64476-64493.
2. Shafik, W., Tufail, A., Namoun, A., De Silva, L. C., & Apong, R. A. A. H. M. (2023). A systematic literature review on plant disease detection: Motivations, classification techniques, datasets, challenges, and future trends. Ieee Access, 11, 59174-59203.
3. Manu, Y. M., Gagana, R., Naveen, K. B., Prashantha, K., & Jadhav, A. (2024, July). Intelligent Recognition of Plant Abnormalities Using Deep Learning Method. In 2024 Second International Conference on Advances in Information Technology (ICAIT) (Vol. 1, pp. 1-6). IEEE.
4. Balafas, V., Karantoumanis, E., Louta, M., & Ploskas, N. (2023). Machine learning and deep learning for plant disease classification and detection. IEEE Access.
5. Hosny, K. M., El-Hady, W. M., Samy, F. M., Vrochidou, E., & Papakostas, G. A. (2023). Multi class classification of plant leaf diseases using feature fusion of deep convolutional neural network and local binary pattern. IEEE Access, 11, 62307-62317.
6. Liu, Q., Zuo, S. M., Peng, S., Zhang, H., Peng, Y., Li, W., ... & Kang, H. (2024). Development of Machine Learning Methods for Accurate Prediction of Plant Disease Resistance. Engineering.
7. Bouacida, I., Farou, B., Djakhdjakha, L., Seridi, H., & Kurulay, M. (2024). Innovative deep learning approach for cross-crop plant disease detection: A generalized method for identifying unhealthy leaves. Information Processing in Agriculture.
8. Polly, R., & Devi, E. A. (2024). Semantic segmentation for plant leaf disease classification and damage detection: A deep learning approach. Smart Agricultural Technology, 9, 100526. Department of CSE, GMRIT Page39 AGROGUIDE: A WEB APPLICATION FOR PERSONALIZED FARMING ASSISTANCE 2025
9. Barhate, D., Pathak, S., Singh, B. K., Jain, A., & Dubey, A. K. (2024). A systematic review of Machine Learning and Deep Learning Approaches in Plant Species Detection. Smart Agricultural Technology, 100605.
10. Shafik, W., Tufail, A., Namoun, A., De Silva, L. C., & Apong, R. A. A. H. M. (2023). A systematic literature review on plant disease detection: Motivations, classification techniques, datasets, challenges, and future trends. Ieee Access, 11, 59174-59203.
11. Jackulin, C., & Murugavalli, S. (2022). A comprehensive review on detection of plant disease using machine learning and deep learning approaches. Measurement: Sensors, 24, 100441.
12. A review on machine learning and deep learning image-based plant disease classification for industrial farming systems. Journal Of Industrial Information Integration, 100572.
13. Durai, S. K. S., & Shamili, M. D. (2024). Smart farming using machine learning and deep learning techniques. Decision Analytics Journal, 3, 100041.
14. Wang, D., Thunéll, S., Lindberg, U., Jiang, L., Trygg, J., Tysklind, M., & Souihi, N. (2024). A machine learning framework to improve effluent quality control in wastewater treatment plants. Science of the total environment,784, 147138.
15. A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools. Smart Agricultural Technology, 3, 100083.