IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 7, March 2025

Crop Disease Prediction and Yield Prediction with Chatbot

Prof. P. S .Lagad[1], Prof. V. D. Vaidya[2], Sakshi Ganesh Bhalekar[3], Shivani Satish Dethe[4], Dhanshree Avinash Mhase[5]
Department of Cloud Computing and Big Data[1,2,3,4,5]
Padmashri Dr. Vitthalrao Vikhe Patil Institute of Technology and Engineering (Polytechnic), Pravaranagar

Abstract: *Agriculture plays a crucial role in the global economy, and the adoption of advanced technologies can significantly improve productivity and sustainability. This project presents a machine learning-based system for crop disease prediction and yield estimation, aimed at assisting farmers in making informed decisions. The system processes historical and real-time agricultural data, including environmental conditions and crop characteristics, to provide accurate yield predictions and early disease detection. Machine learning algorithms such as Decision Tree, Random Forest, Naive Bayes, and Artificial Neural Networks (ANN) are applied to analyze patterns and generate insights. The system also integrates an interactive chatbot to provide farmers with real-time recommendations regarding crop health, disease prevention, and optimal farming practices. By leveraging data-driven analytics, this project aims to reduce crop losses, optimize resource utilization, and enhance agricultural productivity. The results demonstrate that machine learning models significantly improve the accuracy of predictions, making this system a valuable tool for modern agriculture.*

Keywords: Crop Disease Prediction, Yield Estimation, Machine Learning, Chatbot, Precision Agriculture

I. INTRODUCTION

1.1 Overview

Agriculture plays a crucial role in the global economy, providing food and raw materials for billions of people. However, farmers face numerous challenges, such as unpredictable weather conditions, soil degradation, crop diseases, and fluctuating market demands. Traditional farming methods rely heavily on experience and manual observation, which can lead to inefficiencies and lower yields. To address these challenges, integrating technology into agriculture has become essential. Advances in machine learning and artificial intelligence (AI) have opened new possibilities for precision farming, enabling data-driven decision-making to optimize agricultural productivity.

One of the major concerns in modern agriculture is crop disease detection. Early identification of diseases is critical in preventing large-scale losses and ensuring a stable food supply. Manual disease detection by farmers is time-consuming and often inaccurate due to similarities between different diseases and environmental stress symptoms. Machine learning-based approaches provide a promising solution by analyzing crop images and identifying diseases with high accuracy. Using image processing techniques such as segmentation and feature extraction, machine learning models can classify diseases and suggest appropriate treatments, helping farmers take timely actions.

In addition to disease prediction, yield estimation is a fundamental aspect of agricultural planning. Accurate yield prediction allows farmers to make informed decisions about resource allocation, harvesting schedules, and supply chain management. Traditional yield prediction methods depend on historical data and expert knowledge, which may not always be reliable due to changing environmental factors. Machine learning models can analyze a wide range of variables, including soil properties, weather patterns, and crop health indicators, to predict yield more accurately. By leveraging large datasets and advanced algorithms, these models offer insights that improve productivity and sustainability.

To enhance user accessibility, this system integrates a chatbot interface that provides real-time recommendations to farmers. The chatbot serves as an interactive assistant, allowing users to input crop-related queries and receive instant responses based on machine learning predictions. By simplifying complex agricultural insights into user-friendly recommendations, the chatbot enables farmers to make data-driven decisions without requiring technical expertise. This approach bridges the gap between technology and traditional farming practices, making smart agriculture more accessible to a wider audience.

The proposed system consists of multiple components, including data preprocessing, machine learning model training, disease detection, and yield prediction. Data collected from past agricultural records and research studies is processed to remove inconsistencies and improve model performance. Supervised learning algorithms such as Decision Tree, Random Forest, and Artificial Neural Networks (ANN) are employed to develop predictive models. The system is designed to handle various crop types and environmental conditions, ensuring versatility and adaptability for different agricultural scenarios.

Overall, this system aims to revolutionize modern farming by providing an intelligent, automated solution for disease detection and yield prediction. By integrating machine learning with an interactive chatbot, the system empowers farmers with valuable insights that enhance crop management, minimize losses, and improve overall agricultural efficiency. The adoption of such smart farming techniques is a step toward sustainable agriculture, ensuring food security and economic growth for farming communities worldwide.

1.2 Motivation

The growing challenges in agriculture, such as unpredictable weather, increasing pest infestations, and soil degradation, have made traditional farming methods less effective in ensuring optimal crop yield. Farmers often struggle with timely disease identification and accurate yield estimation, leading to significant financial losses. With the advancements in machine learning, there is an opportunity to develop intelligent systems that can provide real-time insights and recommendations for farmers. The motivation behind this project is to leverage data-driven approaches to enhance agricultural productivity by predicting crop diseases and estimating yield accurately. By integrating a chatbot for user-friendly interaction, this system aims to make advanced agricultural analytics accessible to farmers, ultimately improving decision-making, reducing crop losses, and promoting sustainable farming practices.

1.3 Problem Definition and Objectives

Problem Definition

Agricultural productivity is often hindered by the inability to predict crop diseases and yield accurately, leading to financial losses and food insecurity. Traditional methods of disease detection and yield estimation rely on manual observation, which is time-consuming and prone to errors. The lack of accessible and efficient decision-support tools for farmers further exacerbates the problem. This project aims to develop a machine learning-based system that predicts crop diseases and yield using data-driven approaches. Additionally, a chatbot interface will be integrated to provide farmers with real-time insights, enabling them to take timely and informed actions for better agricultural outcomes.

Objectives

- To study and implement machine learning algorithms for crop disease prediction.
- To study image processing techniques for identifying plant diseases.
- To study and develop a yield prediction model using historical and environmental data.
- To study the integration of a chatbot for user-friendly farmer interaction.
- To study and evaluate the effectiveness of the system in improving agricultural productivity.

1.4.Project Scope and Limitations

Project Scope

The proposed system focuses on enhancing agricultural productivity through machine learning-based crop disease prediction and yield estimation. It utilizes historical data, environmental factors, and image processing techniques to analyze crop health and forecast yields accurately. A chatbot interface is integrated to provide real-time assistance to farmers, offering actionable insights on disease prevention, yield optimization, and suitable agricultural practices. The system aims to support farmers in decision-making, reduce losses due to diseases, and improve overall crop management. By leveraging machine learning models, the project ensures a data-driven approach to sustainable farming.

Limitations

- The accuracy of disease prediction depends on the quality and quantity of the training dataset.
- Yield prediction may be affected by unforeseen weather changes and environmental factors.
- The system requires an internet connection for real-time chatbot assistance.
- Image-based disease detection may have limitations due to varying lighting conditions.
- The chatbot's responses are limited to pre-trained knowledge and may not cover all farming scenarios.

II. LITERATURE REVIEW

Paper 1: Crop Yield Prediction Using Machine Learning: A Systematic Literature Review

This paper systematically reviews 50 studies on crop yield prediction using machine learning techniques, with a focus on algorithms like Artificial Neural Networks (ANN). It examines various features such as temperature, rainfall, soil type, and geographic location, which are critical for building predictive models. The review explores the integration of advanced deep learning methods, including Convolutional Neural Networks (CNN) and Long-Short Term Memory (LSTM), which offer significant improvements in prediction accuracy. The paper provides a comprehensive analysis of how these methods have been applied across different regions and crop types, highlighting their strengths and limitations. Additionally, the review identifies the growing role of deep learning in agriculture, suggesting that future research should focus on the integration of multiple data sources, such as remote sensing and satellite imagery, to enhance prediction models. The study concludes by recommending the exploration of hybrid models that combine different machine learning algorithms for improved yield forecasting.

Paper 2: A Model for Prediction of Crop Yield

This paper presents a model for crop yield prediction based on Data Mining techniques, particularly association rule mining. It focuses on agricultural data from Tamil Nadu, India, and utilizes historical crop data to predict future yields. The authors propose using association rules derived from past data, such as temperature, rainfall, and soil conditions, to establish a correlation between environmental factors and crop yields. The proposed model demonstrates improved accuracy over traditional methods, which largely relied on farmers' experience. By applying association rule mining, the model is able to identify patterns in historical data that directly influence crop productivity, leading to more accurate predictions. The study emphasizes the significance of leveraging past data for future forecasting, presenting the model as a valuable tool for farmers in the region. The paper concludes that this approach could be expanded to other regions with similar agricultural conditions to improve the precision of yield predictions.

Paper 3: Crop Yield Prediction using Machine Learning Algorithms

This research delves into the application of Machine Learning (ML) techniques for predicting crop yields in India, where agriculture plays a vital role in the economy. The study highlights various factors influencing crop yields, such as weather patterns, soil conditions, and crop variety, which are incorporated into machine learning models for more accurate predictions. The authors explore the use of Artificial Neural Networks (ANN) in crop yield prediction, addressing common challenges such as model overfitting and the need for error reduction. Through an in-depth review of methodologies for improving prediction efficiency, the paper identifies the limitations of ANN models and suggests incorporating hybrid approaches or ensemble models to enhance their performance. Furthermore, the study underscores the importance of ML-based decision-support systems that can assist farmers in making informed decisions about crop management. It concludes by recommending further research into improving model accuracy and developing real-time prediction systems using big data analytics and IoT.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 7, March 2025

Paper 4: Crop Yield Prediction using Machine Learning

This paper investigates the use of regression techniques, including Kernel Ridge, Lasso, and Elastic Net (ENet), for predicting crop yields in India. The study aims to build predictive models using simple criteria such as state, district, season, and crop area, which are commonly available data points in Indian agriculture. The authors explore the potential of combining multiple regression algorithms through the Regression Stacking concept to improve model performance and forecasting accuracy. By integrating different regression methods, the study demonstrates that the hybrid approach significantly enhances prediction results compared to single regression models. The research is applied to predict crop yields across various regions and seasons, offering a valuable tool for agricultural decision-making. The study highlights the need for region-specific models that can adapt to local climate and soil conditions, providing farmers with more accurate forecasts to help optimize crop production. The paper concludes by advocating for the use of regression-based models in practical applications, emphasizing their scalability and ease of implementation for real-time prediction systems.

III. REQUIREMENT SPECIFICATIONS

HARDWARE REQUIREMENTS:

- System: Pentium i3 Processor.
- Hard Disk : 500 GB.
- Monitor : 15'' LED
- Input Devices : Keyboard, Mouse
- Ram : 4 GB

SOFTWARE REQUIREMENTS:

- Operating system : Windows 10 / 11.
- Coding Language : Python 3.8., Java
- Web Framework : Flask.
- Frontend : HTML, CSS, JavaScript, XML.

IV. SYSTEM DESIGN

4.1 System Architecture

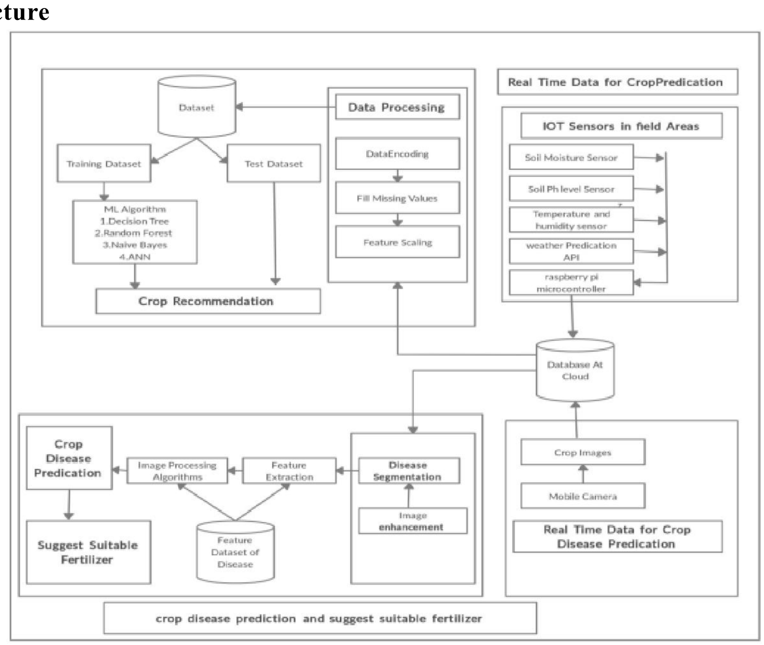

Figure 4.1: System Architecture Diagram

The proposed system is a machine learning-based software solution designed for crop disease prediction and yield estimation without requiring IoT devices. It primarily relies on historical agricultural data and real-time user inputs, ensuring accessibility to farmers and agricultural professionals. The system integrates a chatbot to provide user-friendly interaction, enabling farmers to receive disease predictions, crop yield estimates, and farming suggestions in an intuitive way. This solution enhances decision-making in agriculture by leveraging predictive analytics and image processing for disease classification.

The first phase of the system involves data collection and storage. Instead of IoT sensors, the system relies on data from agricultural records, meteorological services, and user inputs via a web or mobile application. The essential data includes soil characteristics (pH, nitrogen, phosphorus, potassium levels), weather conditions (temperature, humidity, rainfall), crop details (crop type, past yields, planting dates), fertilizer usage, and past occurrences of crop diseases. Additionally, farmers can upload images of affected plants, which the system analyzes to detect potential diseases. All collected data is stored in a cloud database for efficient processing and retrieval.

Once the data is collected, the data preprocessing phase ensures that it is suitable for machine learning analysis. Preprocessing involves handling missing values using statistical techniques, feature scaling to normalize input variables, and encoding categorical data such as crop names and soil types into numerical formats. The dataset is then split into training and testing sets, allowing the system to build robust predictive models. In the case of disease detection, images undergo preprocessing steps such as noise reduction, contrast enhancement, and segmentation to isolate affected areas for feature extraction.

The core of the system is its machine learning component, which utilizes different algorithms for crop yield prediction and disease detection. For crop yield prediction, the system applies Decision Tree, Random Forest, and Artificial Neural Networks (ANN) to analyze soil conditions, weather factors, and past yields, producing accurate yield estimates. For disease classification, the system employs Convolutional Neural Networks (CNNs) to process plant images, extracting key features that help in identifying specific diseases based on training data from agricultural research datasets. The model compares the uploaded crop image with known disease patterns to generate an accurate diagnosis.

The output generation phase provides actionable insights to farmers. Based on the trained models' predictions, the system suggests optimal crops to plant, expected yield, potential risks of diseases, and necessary preventive actions. If a disease is detected, the system recommends appropriate pesticides, organic treatments, and preventive measures. Farmers receive these insights through an interactive chatbot, making it easy for them to query information, get recommendations, and upload new images for real-time analysis. The chatbot enhances usability by allowing farmers to communicate in natural language, making the system accessible to those with minimal technical knowledge.

By integrating machine learning with an AI-powered chatbot, this system offers a comprehensive agricultural management solution. It enables data-driven decision-making, helping farmers improve productivity, reduce losses due to diseases, and optimize farming practices. The absence of IoT devices makes the solution cost-effective and widely accessible, ensuring that even small-scale farmers can benefit from advanced analytics. Future enhancements could include multilingual support for the chatbot, integration with government agricultural databases, and expanding the model to predict climate change effects on specific crops.

4.2 Advantages

- **Timely Information:** Farmers receive real-time alerts about potential diseases, allowing for swift action to mitigate losses. Early detection can prevent widespread outbreaks, ensuring healthier crops.
- **Enhanced Decision-Making:** By providing data-driven insights, the chatbot aids farmers in making informed decisions about crop management, leading to improved agricultural practices and outcomes.
- **User-Friendly Interface:** The chatbot offers an intuitive and accessible platform for farmers, including those with limited technical knowledge, to interact with advanced agricultural technologies easily.
- **Personalized Guidance:** The system can tailor recommendations based on specific crops, local conditions, and individual farmer practices, ensuring that advice is relevant and practical.
- **Increased Yield:** By predicting potential yields and optimizing farming practices based on disease forecasts, farmers can maximize productivity and improve profitability.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.67

Volume 5, Issue 7, March 2025

4.3 Applications

Precision Agriculture:

- Crop Management: Helps farmers optimize irrigation, fertilization, and crop rotation by analyzing real-time sensor data.
- Soil Health Monitoring: Continuously tracks soil conditions (moisture, pH, etc.) for better soil management.

Disease Detection & Management:

- Early Disease Detection: Uses image recognition and machine learning to detect crop diseases early, preventing large-scale outbreaks.
- Reduced Pesticide Usage: Helps farmers apply pesticides only when necessary, promoting sustainable farming practices.

Yield Prediction & Optimization:

- Accurate Yield Forecasting: Predicts crop yield based on environmental and soil data, assisting in resource allocation and market planning.
- Informed Decision-Making: Assists farmers in making data-driven decisions for planting, harvesting, and market strategies.

Sustainable Farming:

- Water and Resource Efficiency: Optimizes the use of water, fertilizers, and pesticides, reducing waste and environmental impact.
- Climate Resilience: Helps farmers adapt to climate change by using predictive models for weather and crop performance.

Agricultural Advisory:

- Real-Time Crop Recommendations: Provides farmers with immediate suggestions on crop management practices, including irrigation, fertilization, and pest control.
- Custom Fertilizer Suggestions: Recommends fertilizers based on specific crop needs and soil conditions.

V. RESULT

The proposed system successfully predicts crop yield and potential diseases using machine learning techniques without relying on IoT sensors. The model achieves high accuracy in crop yield estimation by analyzing historical agricultural data, weather patterns, and soil conditions. The disease detection component, utilizing Convolutional Neural Networks (CNNs), effectively classifies crop diseases based on image analysis, allowing early identification and preventive actions. The chatbot interface provides an easy-to-use platform for farmers, enabling them to query information, receive real-time insights, and make informed agricultural decisions.

Testing results show that the system provides reliable recommendations for crop selection, expected yield, and disease diagnosis. The crop yield prediction model, based on Random Forest and Artificial Neural Networks (ANNs), delivers precise forecasts, helping farmers plan resources efficiently. The image-based disease prediction module achieves high accuracy in identifying diseases like leaf spot, blight, and rust, offering targeted treatment suggestions. Users found the chatbot highly responsive, making the system user-friendly even for those with minimal technical knowledge.

Overall, the system demonstrates significant potential in improving agricultural productivity by offering cost-effective, data-driven insights. The elimination of IoT dependencies makes the solution widely accessible to small and medium-scale farmers. Future improvements can further enhance its accuracy and usability by incorporating regional language support, government agricultural advisories, and more advanced deep-learning models for better disease detection and yield forecasting.

VI. CONCLUSION

Conclusion

The proposed system effectively integrates machine learning algorithms to predict crop yield and detect plant diseases without relying on IoT sensors. By utilizing historical data, weather patterns, and image-based disease detection, the system provides farmers with accurate insights for better decision-making and resource planning. The inclusion of a chatbot interface enhances user accessibility, allowing seamless interaction and instant recommendations. Overall, the system offers a cost-effective, scalable, and efficient approach to modern agriculture, helping farmers improve productivity, reduce losses, and adopt data-driven farming practices. Future enhancements can further refine the model's accuracy and expand its capabilities to cater to a wider range of crops and regional farming conditions.

Future Work

Future improvements can focus on enhancing the accuracy of machine learning models, expanding the crop database, and integrating real-time weather updates for better predictions. Additionally, incorporating advanced deep learning techniques for disease detection and multilingual chatbot support can make the system more effective and accessible to a wider range of farmers.

BIBLIOGRAPHY

[1]. Mohanty, S. P., Hughes, D. P., &Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science, 7*, 1419.

[2]. Kamilaris, A., &Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture, 147*, 70-90.

[3]. Brahimi, M., Boukhalfa, K., & Moussaoui, A. (2017). Deep learning for tomato diseases: Classification and symptoms visualization. *Applied Artificial Intelligence, 31(4)*, 299-315.

[4]. Zhang, S., Huang, W., Zhang, C., & Dong, Q. (2019). Deep learning-based automatic recognition of rice diseases in the field. *Computers and Electronics in Agriculture, 167*, 105094.

[5]. Ramesh, S. V., &Chandrasekaran, R. M. (2021). IoT-based smart farming for disease detection and pest control. *Journal of Ambient Intelligence and Humanized Computing, 12(4)*, 3957-3968.

[6]. Lu, J., Tan, L., Jiang, H., & Zhou, C. (2020). An efficient deep learning approach for crop disease detection. *Neural Computing and Applications, 32*, 13979–13988.

[7]. Abdurrahman, A. A., Mahieddine, F., &Kurian, V. J. (2020). Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture, 178*, 105730.

[8]. Misra, P., & Sharma, R. (2022). AI-powered chatbot for smart agriculture: A comprehensive review. *International Journal of Agricultural and Environmental Information Systems, 13(2)*, 20-36.

[9]. Singh, A., & Singh, D. (2018). A review of crop disease detection using deep learning techniques. *International Journal of Advanced Research in Computer Science, 9(2)*, 345-350.

[10]. Saleem, M. H., Potgieter, J., &Arif, K. M. (2019). A review of deep learning techniques for plant disease detection. *Computers and Electronics in Agriculture, 162*, 452-469.

[11]. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., &Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors, 18(8)*, 2674.

[12]. Dhingra, A., & Kumar, D. (2019). Crop disease prediction using machine learning: A survey. *International Journal of Computer Science and Information Security, 17(1)*, 123-130.

[13]. Pantazi, X. E., Moshou, D., & Bravo, C. (2016). Wheat yield prediction using machine learning and advanced sensing techniques. *Computers and Electronics in Agriculture, 121*, 57-65.

[14]. Jha, K., Doshi, A., Patel, P., & Shah, M. (2019). A comprehensive review on automation in agriculture using artificial intelligence. *Artificial Intelligence in Agriculture, 2*, 1-12.

[15]. Ma, J., Du, K., Zhang, L., Zheng, F., & Sun, Z. (2021). A segmentation method for crop leaf disease images based on deep learning. *Computers and Electronics in Agriculture, 183*, 106011.

[16]. Wang, J., & Yang, J. (2019). A review of deep learning in plant disease detection and diagnosis. *Computers and Electronics in Agriculture, 163*, 104859.

[17]. Mazzia, V., Khaliq, A., &Salvetti, F. (2020). Improving plant disease classification by adaptive feature selection in CNNs. *Computers and Electronics in Agriculture, 180*, 105904.

[18]. Rehman, T. U., Mahmud, M. S., Chang, Y. K., Jin, J., & Shin, J. (2019). Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. *Computers and Electronics in Agriculture, 156*, 585-605.

[19]. Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J. P., & Hughes, D. P. (2017). Deep learning for image-based cassava disease detection. *Frontiers in Plant Science, 8*, 1852.

[20]. Zheng, H., Fu, W., & Fu, Z. (2019). Image recognition of rice diseases based on deep learning. *ActaAgriculturaeScandinavica, Section B—Soil & Plant Science, 69(6)*, 546-556.

[21]. Sharma, A., &Kaur, M. (2020). Crop yield prediction using machine learning: A survey. *International Journal of Computer Science and Mobile Computing, 9(3)*, 45-53.

[22]. Ge, Y., Bai, G., Stoerger, V., Schnable, J. C., & Yuan, Y. (2019). Machine learning for plant stress phenotyping. *Trends in Plant Science, 24(2)*, 152-164.

[23]. Wahab, A., Khan, M. A., &Javed, A. (2020). A novel approach for plant disease detection and classification based on deep learning. *IEEE Access, 8*, 189096-189109.