International Journal of Engineering Science Invention Research & Development; Vol. 12, Special Issue 11, May 2026
www.ijesird.com, E-ISSN: 2349-6185

AI-DRIVEN CROP YIELD PREDICTION FOR SUSTAINABLE AGRICULTURE

PAVAN U, SUHAS S, SAKESH P, SHASHANK D URS, RAJESH P C
U G Students, K S School of Engineering and Management, Bengaluru
pavanumesh221@gmail.com,srisuhasgowda24@gmail.com,sakesh28sk@gmail.com,shashankdirs77@gmail.com, Rajeshnaidu2848@gmail.com

Abstract- **Agriculture plays a vital role in feeding the global population, and ensuring optimal crop yield is critical for food security. The project focuses on building a machine learning model that recommends a suitable crop to be grown by using user inputs based on environmental factors such as temperature, humidity, rainfall and so on. This practical application of AI in agriculture enables even small-scale farmers to benefit from advanced predictive analytics, leading to improved yields, optimized resource utilization, and more sustainable farming practices.**

Keywords**: Data-Driven Insights, Climate Change, Machine learning, Food security**

1. INTRODUCTION

The objective of this project is to develop a machine learning model that accurately predicts crop yield based on multiple environmental, meteorological, and soil-related factors. The model should be able to analyze historical agricultural data and provide reliable yield predictions to help farmers make informed decisions about crop selection, resource allocation, and risk management. Farmers face challenges in choosing the right crop to grow due to effecting factors like soil type, weather, rainfall, temperature, and market demand. Accurately predicting crop yields is also difficult because of variables such as weather, soil quality, and farming.

The motivation behind this project is to address the challenges of unpredictable crop yields caused by climate change, soil conditions, and fluctuating weather patterns. Traditional yield estimation methods are often inaccurate, leading to financial losses for farmers and food security concerns. By leveraging machine learning, this project aims to provide accurate, data-driven insights to help farmers make informed decisions. Potential applications of this project include:

- Precision Farming – Assisting farmers in optimizing resource allocation such as water, fertilizers, and pesticides, leading to higher productivity and reduced environmental impact.
- Risk Management – Helping farmers plan for adverse weather conditions or climate variations by providing early warnings and adjusting farming strategies accordingly.
- Agricultural Advisory – Provide farmers with actionable insights via apps or web platforms.
- Global Food Security – Enhance crop production forecasts to support food security.

This project can significantly improve crop yield prediction, leading to more efficient farming practices and resource utilization. It helps farmers mitigate risks related to weather and climate change, boosting productivity.

1.1 OBJECTIVES

The primary objectives of this project are:

1. Develop a Machine Learning Model
   - Build a model that accurately predicts crop yield based on key agricultural factors.
2. Data Collection & Preprocessing
   - Gather comprehensive datasets on weather, soil, and crop conditions.
   - Clean and preprocess the data by handling missing values and normalizing features.
3. Real-Time Integration
   - Integrate live weather data from APIs like Open Weather to enhance predictions.
   - Incorporate real-time soil and weather updates to ensure adaptive forecasting.
4. Create a User-Friendly Application
   - Develop a web application using Stream lit to allow easy user interaction.

- Enable farmers to input local conditions and receive real-time crop yield predictions.

5. Provide Actionable Insights

- Offer recommendations on resource allocation like water and fertilizers.
- Help farmers optimize farming practices based on yield predictions.
- 4o mini

2. LITERATURE SURVEY

Recent advancements in crop yield prediction have transitioned from traditional methods to machine learning (ML) and remote sensing technologies. Traditional prediction models often relied on basic statistical methods and expert knowledge, which lacked the flexibility and accuracy needed for modern agricultural environments. Machine learning techniques like regression models, decision trees, and random forests have significantly improved prediction accuracy by analyzing complex datasets with multiple influencing factors.

Deep learning models, such as convolutional neural networks (CNNs), are increasingly being applied to process large and complex datasets, as demonstrated by Li et al. (2021). The incorporation of IoT devices also offers real-time data on soil moisture, temperature, and nutrient levels, improving prediction accuracy. However, challenges remain in ensuring data consistency and the ability of models to generalize across diverse regions and crop types. As the agricultural sector adapts to climate change, models need to account for these unpredictable factors. Future research will likely focus on refining data integration techniques and improving model adaptability.

2.1 EXISTING MODELS, TECHNIQUES, AND METHODOLOGIES

Several techniques have been used in previous crop yield prediction systems:

- Regression Models: Used to predict crop yields based on weather, soil, and irrigation factors.
- Decision Trees: Analyzes data to predict yield by making decisions based on environmental features.
- Neural Networks: Applies deep learning to handle complex datasets and improve yield forecasting accuracy.
- Clustering Algorithms: Groups similar regions or crops for more localized yield predictions.
- Support Vector Machines (SVM): Helps predict yields by finding optimal boundaries in environmental data.
- API-Driven Approaches: Integrates real-time weather and soil data from external APIs to enhance predictions.

3. METHODOLOGY

The study incorporates detailed analysis of the Art of Living International Center as a comparative model of sustainable community development. This case study provides insights into alternative approaches to resource management, community engagement, and integrated sustainability practices that contrast with mainstream smart city implementations

3.1. REGISTRATION

- The registration process allows users to create an account with basic information like name, email, and location
- Preferred crop type (e.g., Cereals, Legumes, Oilseeds)
- Production constraints
- Season crop

3.2 RECOGNITION

- Recognition in the context of a crop yield prediction system typically refers to the process of identifying and verifying the user's identity or data inputs. This can involve:
- User Authentication: Verifying user credentials (e.g., email and password) during login to ensure secure access to the system.
- Crop Recognition: Using image recognition or machine learning models to identify and

classify crops based on images or input data, aiding in accurate yield predictions.

- Environmental Factor Recognition: Analyzing real-time data (weather, soil conditions) to recognize key factors affecting crop yield.

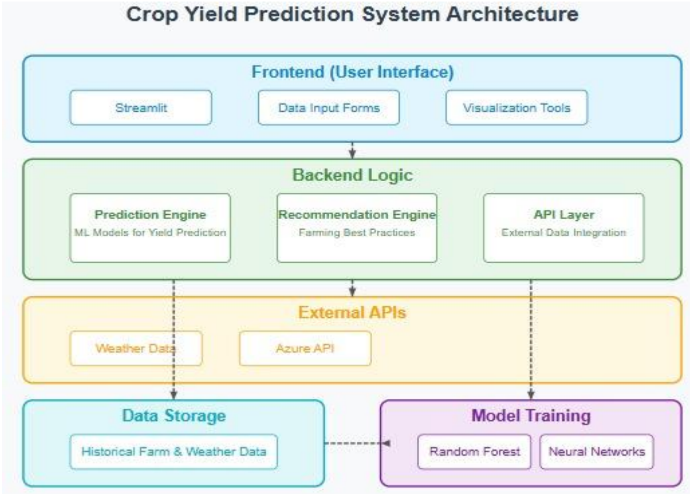
3.3 MODULES USED

1. Data Collection Module:
    - Gathers real-time weather, soil, and satellite data.
    - Integrates external APIs like Open Weather for weather updates.
2. Machine Learning Model Module:
    - Trains models like regression and random forests.
    - Uses environmental data to predict crop yields.
3. Authentication and Registration Module:
    - Manages user sign-up, login, and authentication.
    - Stores user profile data for personalized predictions.
4. Results Visualization Module:
    - A chatbot is being developed to assist users by answering queries related to crop recommendation, production.
5. Prediction Module:
    - Provides crop yield predictions based on input data
    - Uses trained models to forecast future crop performance.

3.3 DATA FLOW DIAGRAM (DFD)

DFD Level 0 - User Interaction:

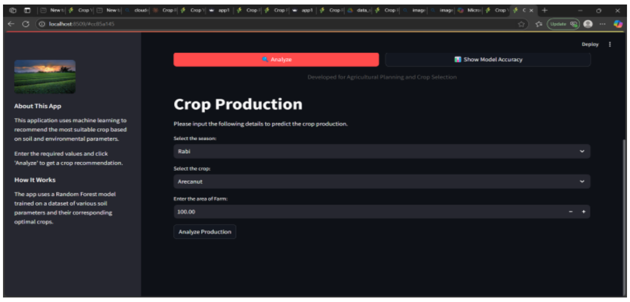
- User inputs crop preferences → System processes data → Recommendations
- DFD Level 1 - Crop Recommendation Module:
- Inputs: User preferences.
- Processes: Machine learning model analyzes data and predicts suitable crop.
- Output: Personalized crop suggestions.
- DFD Level 1 - Real-Time Data Integration:
- Inputs: Weather API.
- Processes: Fetch and process live data.
- Outputs: Display weather conditions.

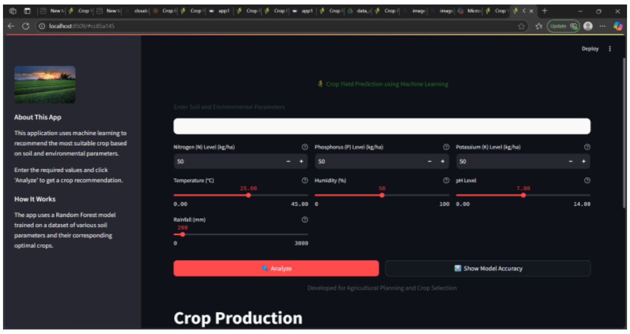


3.4 REQUIREMENT SPECIFICATION

Hardware Requirements:

- A system with a minimum of 4GB RAM and i3 processor (or equivalent).
- Stable internet connection for API data fetching.
- Software Requirements:
- Python 3.x
- stream lit
- Pandas, NumPy, Scikit-learn
- Google Maps API, Open Weather API
- Pre-trained ML model (model.pkl)





4. IMPLEMENTATION AND RESULT

4.1 IMPLEMENTATION

The Data Collection phase involves gathering historical weather, soil, and crop data from various sources like APIs, sensors, and satellite imagery. This data is then processed and cleaned in the Data Preprocessing phase to ensure it is ready for use in machine learning models The main implementation components include:

- Model Training: Applying machine learning algorithms like regression, random forests, and neural networks to train the model.
- Model Training: Applying machine learning algorithms like regression, random forests, and neural networks to train the model.
- User Interface: Developing an intuitive web interface for users to input data and view predictions.
- Real-Time Data Integration: Fetching live weather and environmental data to adjust predictions dynamically.
- Chatbot Development (In Progress): A chatbot feature is being developed to enhance user interaction for farmer related queries.

4.2 RESULTS

4.2.1 Crop Recommendations

The system provides reliable crop yield forecasts based on real-time and historical data, improving planning and decision-making for farmers. Farmers can view detailed predictions, including the impact of weather and environmental factors on crop yields, helping them make better crop management decisions. Supports decision-making that enhances crop productivity and sustainability.

4.2.2 Real-Time Data Accuracy

- Real-time data accuracy is crucial for the effectiveness of the crop yield prediction system, as it directly impacts the precision of predictions.
- The system relies on live weather updates, soil conditions, and environmental factors to adjust forecasts.

4.2.3 User Experience

- Users found the interface simple and interactive, enhancing the overall planning experience.
- The clear presentation of crop details helped users make more informed decisions.

4.3 LIMITATIONS

- Chatbot functionality is still under development and not yet integrated.
- Factors like pests, diseases, and unforeseen climate events may not be fully accounted for in the model, affecting prediction accuracy.
- Farmers with limited technical knowledge may find it challenging to interpret the predictions or fully utilize the system’s features.

5. DISCUSSION AND CONCLUSION

5.1 KEY FINDINGS

Improved Prediction Accuracy: Integrating real-time weather, soil, and environmental data significantly enhances the accuracy of crop yield forecasts. Impact of Local Conditions: Localized data, such as soil type and irrigation methods, plays a critical role in making more precise predictions. Climate Variability: The system effectively adapts to changing climatic conditions, improving predictions even during periods of unusual weather patterns. Resource Optimization.

The model helps optimize resource allocation (water, fertilizers), leading to more efficient farming practices. User Engagement: Farmers benefit from actionable insights, making it easier to plan and make informed decisions about crop management.

5.2 FUTURE WORK

Real-Time Field Monitoring: Implementing IoT devices and sensors for real-time field data collection, improving prediction accuracy and enabling better crop management. Integration of More Data Sources: Incorporating additional data such as pest outbreaks, disease reports, and market trends to provide a more holistic prediction. Mobile Application: Developing a mobile version of the system to allow farmers in remote areas with limited internet access to easily access predictions and updates. Climate Change Adaptation: Enhancing the system to better predict crop yields under varying climate change scenarios, ensuring its relevance in future environmental conditions. Farmer Education and Training: Offering educational resources and training programs for farmers to better understand and utilize the system's features for improved decision-making. Mobile Enhanced Model Accuracy: Improving machine learning algorithms to account for more complex factors and improve prediction accuracy across diverse regions and crops.

5.3 CONCLUSION

In conclusion, the crop yield prediction system leverages advanced machine learning models and real-time environmental data to provide accurate and actionable predictions for farmers. By integrating weather, soil, and crop-specific data, the system helps optimize resource usage, improve productivity, and enhance decision-making.

Machine learning helps farmers predict crop yields more accurately using weather, and crop data. It improves decision-making, reduces waste, and increases productivity. With ML, farming becomes smarter, more efficient, and ready for future challenges like climate change.

In summary, machine learning gives farmers the tools to predict and manage their crops more effectively, leading to better harvests and a more sustainable agricultural future. As this technology improves, it will be crucial in ensuring food security and supporting the growth of farming around the world.

REFERENCES

1) Ahamed, A.T.M.S., Mahmood, N.T., Hossain, N., Kabir, M.T., Das, K., Rahman, F., Rahman, R.M., 2015. Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh.
2) Bargoti, S., Underwood, J.P., 2017. Image segmentation for fruit detection and yield estimation in apple orchards. J. Field Rob. 34 (6), 1039–1060. https://doi.org/10. 1002/rob.21699
3) Cheng, H., Damerow, L., Sun, Y., Blanke, M., 2017. Early yield prediction using image analysis of apple fruit and tree canopy features with neural networks. J. Imag. 3 (1), 6. https://doi.org/10.3390/jimaging3010006.
4) Mola-Yudego, B., Rahlf, J., Astrup, R., Dimitriou, I., 2016. Spatial yield estimates of fastgrowing willow plantations for energy based on climatic variables in northern Europe. GCB Bioenergy 8 (6),1093–1105. https://doi.org/10.1111/gcbb.12332
5) Ranjan, A.K., Parida, B.R., 2019. Paddy acreage mapping and yield prediction using sentinel-based optical and SAR data in Sahibganj district, Jharkhand (India). Spatial Inf. Res. https://doi.org/10.1007/s41324-019-00246-4.
6) Ruder, S., 2017. An overview of multi-task learning in deep neural networks. arXiv preprint arXiv:1706.05098.
7) Sujatha, R., Isakki, P., 2016. A study on crop yield forecasting using classification techniques. In: 2016 International Conference on Computing Technologies and Intelligent Data Engineering, ICCTIDE(2016). https://doi.org/10.1109/ICCTIDE.2016.7725357.