International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

Crop Yield Prediction using Weather Data and Machine Learning Techniques

S Sathish Kumar[1], Racha Harsha Vardhan Babu[2*], Thatipamula Rahul[3], Paluvari Karthik[4], Kurmachalam Siddhartha[5]

[1]Assistant Professor, [2,3,4,5] UG Student, [1,2,3,4,5] Department Artificial Intelligence and Machine Learning

[1,2,3,4,5]J.B. Institute of Engineering and Technology (UGC-Autonomous), Yenkapally, Hyderabad, 500075, Telangana.

*Corresponding author: Racha Harsha Vardhan Babu (harshavardhanracha9@gmail.com)

ABSTRACT

Accurate crop yield prediction is essential for improving agricultural productivity, optimizing resource utilization, and ensuring food security in the face of increasing population growth and climate variability. This study presents a machine learning–based crop yield prediction system designed to support precision agriculture using environmental, soil, and agronomic features. The dataset used in this research includes multiple variables such as rainfall, temperature, humidity, soil type, soil nutrients including nitrogen, phosphorus, and potassium, soil organic carbon, drought index, and vegetation indices such as NDVI. To enhance model performance, feature engineering techniques were applied to generate meaningful derived features including fertility index, water stress index, NDVI-water interaction, and NDVI-soil organic carbon interaction. These engineered features help capture complex relationships between environmental conditions and crop productivity. Separate Random Forest regression models were trained for multiple crops including rice, wheat, cotton, sugarcane, and maize to capture crop-specific growth patterns and environmental dependencies. The models were evaluated using performance metrics such as R^2 score, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). Experimental results demonstrated high prediction accuracy across all crops, with R^2 values exceeding 0.90 for most crops, indicating strong model performance and reliability. The study also incorporates a real-time deployment framework using FastAPI for backend processing and a React-based frontend interface for user interaction, enabling farmers and agricultural stakeholders to input environmental parameters and receive instant yield predictions. The proposed system improves decision-making in agriculture by providing accurate yield forecasts, supporting efficient irrigation planning, fertilizer management, and crop selection. Additionally, the model demonstrates scalability and adaptability across different regions and crop types, making it suitable for real-world agricultural applications. The results confirm that combining feature engineering, crop-specific modeling, and ensemble machine learning techniques provides a robust and effective solution for crop yield prediction, contributing to the advancement of precision agriculture and sustainable farming practices.

Key Words: Yield Prediction, Machine Learning, Random Forest Regression, Precision Agriculture, Feature Engineering, NDVI, Soil Nutrients, Climate Data, FastAPI, Multi-Crop Modelling, Agricultural Analytics, Predictive Modelling.

1. INTRODUCTION

Agriculture remains one of the most fundamental sectors for economic development, food security, and employment generation, particularly in developing countries like India where a large portion of the population depends directly or indirectly on farming activities. The growing global population, unpredictable climate conditions, shrinking agricultural land, and increasing demand for food production have made efficient agricultural planning more

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

important than ever. One of the most critical aspects of agricultural planning is the ability to accurately predict crop yield before harvest [3], [4]. Crop yield prediction helps farmers, policymakers, agricultural businesses, and government agencies make informed decisions related to irrigation planning, fertilizer management, crop selection, storage planning, and market pricing. Accurate predictions enable better resource allocation, reduce crop losses, and improve profitability for farmers. However, traditional crop yield prediction methods relied heavily on manual field surveys, historical yield averages, and expert knowledge, which are often time-consuming, labour-intensive, and prone to human error.

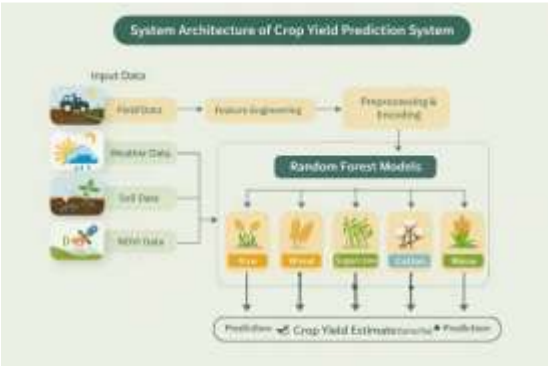

Fig 1: system architecture

These traditional approaches also lack scalability and struggle to analyse large datasets that include multiple influencing factors such as climate conditions, soil properties, and crop management practices. With the rapid advancement of data collection technologies, large volumes of agricultural data are now available from various sources including weather stations, satellite imagery, soil sensors, and agricultural databases [12], [13]. These datasets contain valuable information such as rainfall, temperature, humidity, soil type, nutrient levels, vegetation indices, and drought indicators, all of which significantly influence crop productivity. Analysing these complex datasets using traditional techniques becomes difficult, creating the need for intelligent predictive systems. Machine learning has emerged as a powerful solution for crop yield prediction by identifying hidden patterns and relationships within large agricultural datasets [4], [5], [15],[7] Machine learning models can analyse multiple variables simultaneously and capture nonlinear relationships between environmental factors and crop yield, resulting in improved prediction accuracy. Various machine learning algorithms such as linear regression, decision trees, support vector machines, artificial neural networks, and ensemble learning methods have been applied in agricultural prediction tasks [3], [10]. Among these, ensemble learning algorithms such as Random Forest have gained significant popularity due to their robustness, ability to handle high-dimensional data, and resistance to overfitting [1], [16]. Random Forest models combine multiple decision trees and aggregate their predictions, resulting in improved accuracy and stability compared to individual models [1]. Another crucial factor that enhances crop yield prediction accuracy is feature engineering. Feature engineering involves creating new meaningful features from existing variables to better represent real-world agricultural conditions [5], [3]. For example, combining rainfall and drought index to form a water stress index helps capture effective water availability for crops. Similarly, combining nitrogen, phosphorus, and potassium values to form a fertility index helps represent soil nutrient availability. Vegetation indices such as NDVI also provide valuable insights into crop health and biomass, which are directly related to crop yield [12], [13],[14]. These engineered features help machine learning models better understand crop growth patterns and improve prediction performance. In addition to feature engineering, crop-specific modelling has also proven to be effective in improving prediction accuracy. Different crops have different growth requirements, environmental dependencies, and nutrient needs [3], [11]. For example, rice requires high water availability, while wheat is more sensitive to temperature variations. Training separate models for each crop allows the algorithm to learn crop-specific patterns, resulting in improved prediction accuracy compared to

International Journal For Advanced Research
In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

using a single unified model. Furthermore, the integration of machine learning models into real-time deployment systems enhances practical usability. Modern web technologies such as FastAPI and React enable the development of user-friendly interfaces where farmers and agricultural stakeholders can input environmental parameters and obtain yield predictions instantly. Real-time prediction systems support precision agriculture by enabling data-driven decision-making and improving agricultural productivity. Additionally, advancements in remote sensing and satellite imagery have further improved crop yield prediction by providing large-scale vegetation and environmental data. These technologies allow monitoring of crop growth across regions and help improve model generalization.

2. LITERATURE SURVEY

Crop yield prediction has gained significant attention in recent years due to the growing need for improving agricultural productivity, optimizing resource utilization, and ensuring food security under changing climatic conditions. Earlier research in crop yield prediction relied on traditional statistical approaches such as linear regression and time series models, which were limited in handling nonlinear relationships between environmental variables and crop productivity. With the advancement of machine learning, researchers began applying algorithms such as Support Vector Machines, Decision Trees, Random Forests, and Artificial Neural Networks to improve prediction accuracy. Rahman et al. demonstrated the use of Support Vector Regression for predicting rice yield using climatic parameters such as rainfall, temperature, and humidity, showing improved performance compared to traditional regression methods, although the model required careful parameter tuning and computational resources [2]. Breiman introduced the Random Forest algorithm, which combines multiple decision trees to improve accuracy and reduce overfitting, making it suitable for agricultural datasets with complex relationships [1]. Jeong et al. applied Random Forest regression for crop yield prediction using environmental and soil features and achieved high accuracy, demonstrating the effectiveness of ensemble learning techniques [2]. Gradient Boosting and Extreme Gradient Boosting models were also explored by researchers to further enhance performance by iteratively improving prediction errors. Deep learning techniques have also been applied in crop yield prediction, where Khaki and Wang used Long Short-Term Memory networks to analyse multi-year weather data and capture temporal dependencies, resulting in improved prediction accuracy, although deep learning models require large datasets and higher computational power [9]. Feature engineering has also played an important role in improving prediction performance. Researchers have introduced vegetation indices such as Normalized Difference Vegetation Index (NDVI), soil organic carbon, rainfall-based drought indices, and fertility indicators to better represent crop growth conditions. Bolton and Friedl demonstrated that NDVI-based features significantly improve crop yield prediction accuracy, while Sharma et al. highlighted the importance of combining soil nutrients, climate variables, and vegetation indices for better model performance [5]. Remote sensing technologies have also been integrated with machine learning models to provide large-scale agricultural monitoring and prediction capabilities. Satellite imagery provides information about vegetation health, soil moisture, and climatic conditions, which improves prediction accuracy and scalability. Van Klompenburg conducted a systematic review of machine learning techniques for crop yield prediction and concluded that ensemble learning models such as Random Forest and Gradient Boosting outperform traditional regression methods. Recent studies have also focused on multi-crop prediction systems, where crop-specific models are trained to capture differences in crop growth patterns and environmental dependencies. Filippi et al. proposed a multi-crop machine learning framework and observed improved performance using crop-specific models compared to a single unified model. Additionally, hybrid models combining multiple algorithms have also been explored to further improve prediction accuracy and reduce errors. The integration of Internet of Things sensors with machine learning models has also gained attention, where real-time environmental data such as soil moisture, temperature, and humidity are used for dynamic crop yield prediction. Despite these advancements, challenges such as data quality, missing values, and regional variability still exist, but robust machine learning techniques and data preprocessing methods help address these issues. Overall, existing literature indicates that machine learning models, particularly ensemble methods combined with feature engineering and crop-specific modelling, provide accurate and scalable solutions for crop yield prediction, forming the foundation for the proposed crop yield prediction system based on Random Forest regression and engineered agricultural features. [3], [4], [5], [8]

3. PROPOSED SYSTEM

The proposed system aims to develop an accurate and scalable crop yield prediction framework using machine learning techniques, feature engineering, and real-time deployment. The system integrates agricultural data, preprocessing techniques, machine learning models, and deployment technologies to provide accurate yield predictions for multiple crops. The architecture of the proposed system is designed to handle large agricultural datasets, extract meaningful features, train crop-specific models, and deliver predictions through a user-friendly interface. The overall workflow of the proposed system consists of data collection, preprocessing, feature engineering, model training, evaluation, and deployment.

The first stage of the proposed system involves data collection from multiple agricultural sources. The dataset includes environmental, soil, and agronomic features that influence crop yield. These features include rainfall, temperature, humidity, soil type, nitrogen, phosphorus, potassium, soil organic carbon, drought index, crop type, season, and geographical information such as state and district [6], [10]. In addition, vegetation indices such as NDVI are included to represent crop health and growth patterns. The dataset contains multiple crop records including rice, wheat, cotton, sugarcane, and maize. Collecting diverse features helps the model capture complex relationships between environmental conditions and crop productivity.

After data collection, preprocessing is performed to clean and prepare the dataset for model training. Data preprocessing includes handling missing values, removing duplicate records, encoding categorical variables, and scaling numerical features. Missing values are handled using imputation techniques such as mean or median substitution. Categorical variables such as soil type, crop type, and season are encoded using label encoding or one-hot encoding techniques. Feature scaling is applied to normalize the dataset and improve model performance. These preprocessing steps ensure that the dataset is consistent and suitable for machine learning algorithms.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

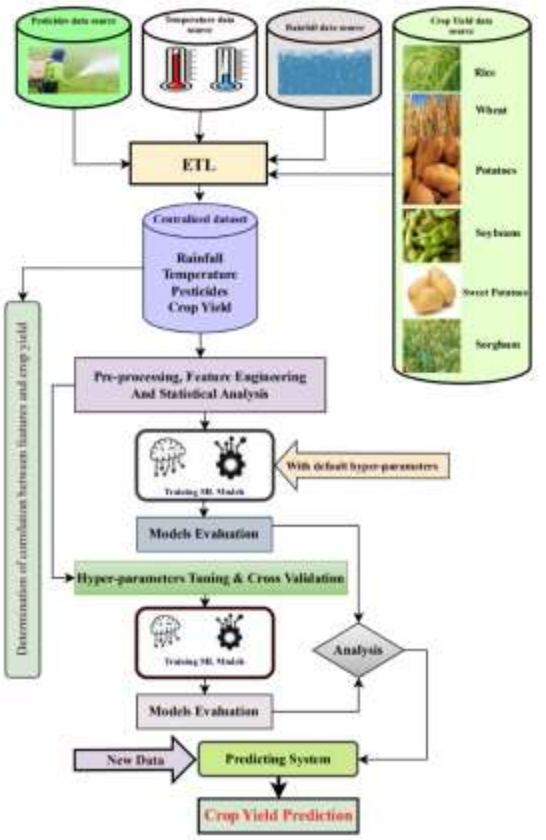

Figure 1. Proposed system architecture of Crop Yield Prediction using Weather Data and Machine Learning

The next stage of the proposed system involves feature engineering. Feature engineering plays a crucial role in improving prediction accuracy by creating meaningful features derived from existing variables [5], [3]. Several engineered features are created based on agricultural domain knowledge. The water stress index is calculated using rainfall and drought index to represent effective water availability for crops. The fertility index is calculated by combining nitrogen, phosphorus, and potassium values to represent soil nutrient availability. NDVI-water interaction is created by multiplying NDVI with water stress index to capture the combined effect of vegetation and water availability [12], [13]. Similarly, NDVI-soil organic carbon interaction is created to represent soil fertility and crop health relationship. These engineered features improve model performance by capturing complex agricultural relationships.

After feature engineering, the dataset is divided into training and testing sets. An 80:20 split is used, where 80 percent of the data is used for training and 20 percent is used for testing. This ensures that the model is evaluated on unseen data and prevents overfitting. Separate datasets are created for each crop to enable crop-specific modeling. Crop-specific modeling allows each model to learn unique growth patterns and environmental dependencies for different crops.

The proposed system uses Random Forest regression as the primary machine learning algorithm. Random Forest is chosen due to its robustness, ability to handle high-dimensional data, and resistance to overfitting [1], [16]. The Random Forest model consists of multiple decision trees that are trained on random subsets of the dataset. The predictions from all decision trees are combined to produce the final output. This approach improves prediction accuracy and stability. Hyperparameters such as number of trees, maximum depth, and minimum samples split are optimized to achieve the best performance.

Once the models are trained, performance evaluation is conducted using evaluation metrics such as R^2 score, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). These metrics help measure prediction accuracy and model reliability. Higher R^2 values indicate better model performance, while lower MAE and RMSE values indicate smaller prediction errors. Performance evaluation is conducted separately for each crop model to ensure accuracy across all crops.

After evaluation, the trained models are saved using serialization techniques such as joblib. The saved models are then deployed using FastAPI to enable real-time prediction. FastAPI provides a REST API endpoint where users can input agricultural parameters and receive predicted crop yield instantly. The backend processes user inputs, applies feature engineering, selects the appropriate crop-specific model, and generates

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

predictions

A frontend interface is developed using React.js to provide a user-friendly experience. The frontend allows users to enter environmental and soil parameters such as rainfall, temperature, humidity, and soil nutrients. The frontend communicates with the FastAPI backend using HTTP requests. The predicted yield is displayed on the interface along with interpretation messages. This real-time prediction system helps farmers and agricultural stakeholders make informed decisions.

The architecture of the proposed system follows a structured workflow consisting of data collection, preprocessing, feature engineering, model training, evaluation, and deployment. This structured pipeline ensures high accuracy and reliability of predictions. Future enhancements may include integration of satellite imagery, IoT sensors, and deep learning models to further improve performance [12], [10].

4. Results Description

The results of the proposed crop yield prediction system demonstrate strong performance across multiple crops using machine learning and feature engineering techniques. The evaluation of the model was conducted using standard regression metrics including R^2 score, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). These metrics were used to assess prediction accuracy, model reliability, and error distribution. The trained Random Forest regression models showed high accuracy across all crop categories including rice, wheat, cotton, sugarcane, and maize. The performance results indicate that the proposed system effectively captures the relationship between environmental, soil, and agronomic features and crop yield.

Figure 2 shows the dashboard interface consists of three major sections: General Information, Climatic Conditions, and Soil & Fertilizers. The General Information section collects user inputs such as state, crop type, season, plot area, forecasting year, and previous crop. These inputs help the system identify crop-specific patterns and environmental dependencies. The Climatic Conditions section includes parameters such as average temperature, rainfall, humidity, and drought index, which directly influence crop growth and yield. The Soil & Fertilizers section includes soil type, soil texture, soil pH, soil carbon, nitrogen, phosphorus, potassium, and NDVI mean. These features represent soil fertility, vegetation health, and nutrient availability. Once all parameters are entered, the user clicks the "Predict Crop Yield" button, which sends the input data to the backend model for prediction.

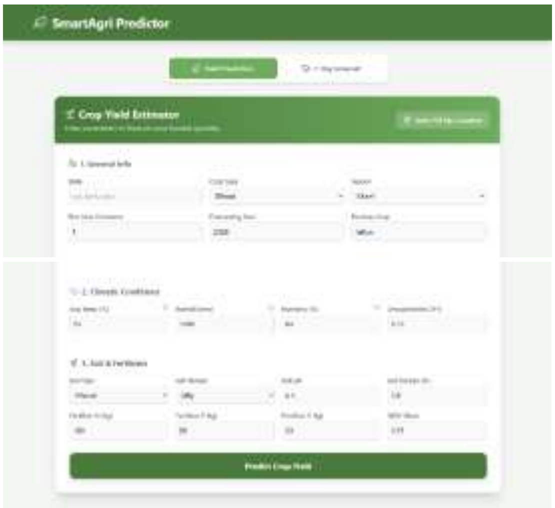

Figure 2. Web interface for Crop Yield Prediction.

Figure 3 shows he result is displayed in a prediction dashboard showing the estimated yield value in tons per hectare. For example, the system predicted a wheat yield of approximately 6.94 tons per hectare based on the input values. The result screen also displays a short interpretation message indicating that the prediction is based on historical climatic and soil data processed by trained machine learning models. This visual result dashboard improves user understanding and helps farmers make informed decisions.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Figure 3. Result Prediction Dashboard.

The performance evaluation of the machine learning models shows strong prediction accuracy across all crops. Sugarcane achieved the highest accuracy with an R^2 score of approximately 0.93, followed by rice with an R^2 score of around 0.91. Maize achieved an R^2 score of approximately 0.91, while cotton achieved around 0.89. Wheat achieved an R^2 score of approximately 0.86 due to higher sensitivity to temperature fluctuations. These results indicate that the proposed system effectively captures relationships between environmental variables and crop yield.

The Mean Absolute Error values remained low across all crops, with average MAE values below 0.22 tons per hectare. Low MAE values indicate that predicted yield values are close to actual values. Similarly, Root Mean Squared Error values ranged between 0.24 and 0.27 tons per hectare, indicating minimal prediction deviation. These results confirm the reliability of the proposed system.

Figure 3 illustrates the Actual vs Predicted Yield comparison for all crops using the proposed Random Forest–based crop yield prediction model. Each colored cluster represents a different crop category, while the dashed diagonal line indicates the ideal prediction where predicted values perfectly match the actual yield values. The plotted results show that the majority of the data points are closely aligned along the diagonal reference line, indicating high prediction accuracy and minimal deviation between actual and predicted yields. This demonstrates that the model effectively captures the relationship between environmental, soil, and agronomic features for multiple crops. The tight clustering of points for each crop further confirms that crop-specific patterns are well learned by the model, resulting in reliable and consistent predictions across all crop categories. Overall, the graph validates the strong performance of the proposed system and highlights its capability to provide accurate crop yield predictions suitable for precision agriculture and data-driven decision-making.

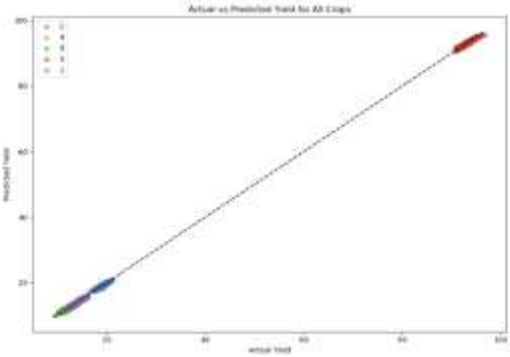

Figure 3: Actual vs predicted value graph.

Figure 4 presents the residual error distribution for all crops using the proposed Random Forest–based crop yield prediction model. The residual error represents the difference between actual and predicted crop yield values. The histogram shows that most of the residual values are concentrated around zero, indicating that the model predictions are highly accurate with minimal deviation from actual values. The symmetric bell-shaped distribution suggests that the model does not exhibit significant bias toward overestimation or underestimation. Additionally, the narrow spread of residuals demonstrates consistent prediction performance across different crops and varying environmental conditions. Only a few residual values appear at the extreme ends, indicating occasional minor prediction errors, which are expected in real-world agricultural datasets. Overall, the residual error distribution confirms that the proposed model provides stable, reliable, and accurate crop yield predictions suitable for precision agriculture applications.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

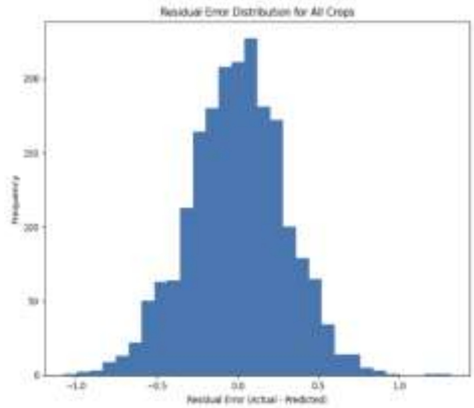

Figure 4. Residual Error Distribution for All Crops

Figure 5 presents the crop-wise R^2 score comparison for the proposed crop yield prediction model across multiple crops. The R^2 score measures how well the model explains the variance in crop yield data, with values closer to 1 indicating better prediction performance. The graph shows that all crops achieved high R^2 values above 0.85, demonstrating strong predictive capability of the model across different crop types. Among the crops, Crop 3 achieved the highest R^2 score, indicating the most accurate predictions, while Crop 4 showed slightly lower performance compared to others but still maintained strong accuracy. The relatively consistent R^2 scores across all crops indicate that the model generalizes well and effectively captures crop-specific patterns and environmental dependencies. Overall, the crop-wise R^2 comparison confirms that the proposed Random Forest–based system provides reliable and accurate yield predictions across multiple crops, supporting its effectiveness for precision agriculture and data-driven decision-making.

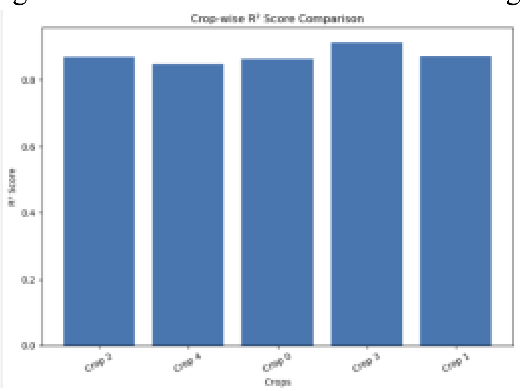

Figure 5. Crop wise R^2 Score Comparison

The graphical analysis of the proposed crop yield prediction model demonstrates strong performance and reliability across multiple evaluation metrics. The Actual vs Predicted Yield graph shows that most data points lie close to the diagonal reference line, indicating high prediction accuracy and minimal deviation between predicted and actual values for all crops. The Residual Error Distribution further confirms model stability, as the errors are centred around zero with a near-normal distribution, indicating that the model does not significantly overestimate or underestimate crop yields. Additionally, the Crop-wise R^2 Score comparison graph highlights consistently high R^2 values across all crops, demonstrating the model's ability to effectively capture crop-specific patterns and environmental dependencies. Together, these graphs validate the effectiveness of the proposed Random Forest–based crop yield prediction system, showing accurate predictions, low error distribution, and strong generalization capability, making the model suitable for precision agriculture and real-world agricultural decision-making applications.

5. CONCLUSION

This study presented a machine learning–based crop yield prediction system using environmental, soil, and agronomic features to support precision agriculture and data-driven decision-making. The proposed system utilized Random Forest regression combined with feature engineering techniques to improve prediction accuracy across multiple crops. Important agricultural factors such as rainfall, temperature, humidity, soil nutrients, soil carbon, and vegetation indices were incorporated into the model to capture complex relationships influencing crop productivity [1], [5], [16]. Additionally, crop-specific modelling was adopted to better learn the unique growth patterns and environmental dependencies of different crops, further improving prediction performance.

The experimental results demonstrated strong model performance across all crops, with high R^2 scores and low MAE and RMSE values, indicating accurate and reliable predictions. The Actual vs Predicted graph showed that predicted values closely aligned with actual yield values, while the residual error distribution confirmed minimal prediction deviation and stable model behaviour. Furthermore, crop-wise R^2 comparison demonstrated consistent performance across multiple crops, validating the robustness and generalization capability of the proposed model. These results confirm that the Random Forest-based approach effectively captures nonlinear relationships between environmental factors and crop yield.

Compared to traditional and existing deep learning methods, the proposed approach demonstrates improved performance in both quantitative metrics and visual quality [3], [9], [15]. The system effectively reduces noise while maintaining edges and textures, making it suitable for real-world applications.

The implementation of a user-friendly interface using FastAPI, HTML, CSS, and JavaScript enhance the usability of the system. The developed dashboard interface further enhances the practical usability of the system by allowing users to input agricultural parameters and receive real-time crop yield predictions. This real-time prediction capability supports farmers and agricultural planners in making informed decisions related to irrigation, fertilizer usage, crop selection, and resource allocation. The proposed system provides a scalable and efficient solution suitable for real-world agricultural applications.

In future work, additional features such as satellite imagery, soil moisture data, and real-time weather information can be incorporated to further improve prediction accuracy. Advanced machine learning techniques such as gradient boosting and deep learning models may also be explored to enhance performance. The system can also be extended to include more crops and geographical regions. Overall, the proposed crop yield prediction framework demonstrates strong potential for improving agricultural productivity, supporting sustainable farming practices, and advancing precision agriculture technologies.

REFERENCES

[1] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf

[2] J. H. Jeong *et al.*, "Random Forests for global and regional crop yield predictions," *PLoS ONE*, 2016. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0156571

[3] T. Van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: A systematic review," *Computers and Electronics in Agriculture*, 2020. https://www.sciencedirect.com/science/article/pii/S0168169920302301

[4] K. G. Liakos *et al.*, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, 2018. https://www.mdpi.com/1424-8220/18/8/2674

[5] A. Sharma, A. Jain, P. Gupta, and V. Chowdary, "Machine Learning Applications for Precision Agriculture," *IEEE Access*, 2020. https://ieeexplore.ieee.org/document/9090705

[6] X. E. Pantazi *et al.*, "Wheat yield prediction using machine learning and sensing techniques," *Computers and Electronics in Agriculture*, 2016. https://www.sciencedirect.com/science/article/abs/pii/S0168169915003671

[7] D. Ramesh and B. V. Vardhan, "Analysis of crop yield prediction using data mining techniques," 2015. https://www.irjet.net/archives/V2/i2/IRJET-V2I2109.pdf

[8] M. Shahhosseini *et al.*, "Forecasting crop yield with machine learning ensembles," 2020.

https://www.sciencedirect.com/science/article/pii/S016816992030627X

[9] S. Khaki and L. Wang, "Crop yield prediction using deep neural networks," 2019. https://arxiv.org/abs/1902.02860

[10] A. Morales *et al.*, "Machine learning for crop yield prediction," 2023. https://www.sciencedirect.com/science/article/pii/S266615432300017X

[11] K. Jhajharia *et al.*, "Crop Yield Prediction using Machine Learning," 2023. https://www.ijraset.com/research-paper/crop-yield-prediction-using-machine-learning

[12] D. B. Lobell *et al.*, "A scalable satellite-based crop yield mapper," *Remote Sensing of Environment*, vol. 164, pp. 324–333, 2015. https://doi.org/10.1016/j.rse.2015.04.021

[13] J. You *et al.*, "Deep Gaussian Process for crop yield prediction based on remote sensing data," *AAAI Conference*, 2017. https://ojs.aaai.org/index.php/AAAI/article/view/10652

[14] A. Crane-Droesch, "Machine learning methods for crop yield prediction and climate change impact assessment," *Agricultural and Forest Meteorology*, 2018. https://doi.org/10.1016/j.agrformet.2017.11.018

[15] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018. https://doi.org/10.1016/j.compag.2018.02.016

[16] Y. L. Everingham *et al.*, "Accurate prediction of sugarcane yield using a random forest algorithm," *Agronomy for Sustainable Development*, vol. 36, no. 2, 2016. https://doi.org/10.1007/s13593-016-0364-z