Automated Crop Yield Prediction With Yield Enhancement Recommendations Using ML

Ch. Phanindra[1], G. Nagalakshmi[2], J. Navya[3], R. Vasantha[4], A. Hemalatha[5]
Department of CSE – Artificial Intelligence
KKR & KSR Institute of Technology and Sciences, Guntur
Email: phanindrakits@gmail.com[1], gudenagalakshmiii@gmail.com[2], 22jr1a4330@gmail.com [3], vasantharamisety@gmail.com [4], amadahemalatha@gmail.com [5]

***Abstract*—The effective planning in agriculture requires**

ac- curate prediction of the crop yield but, most of the machine learning–based techniques to this effect are only concerned with estimating the crop yield and they do not provide actionable strategies for productivity improvement. This makes them less useful for farmers. To fill this hole in knowledge, a crop yield prediction system that combines the yield forecasting and yield enhancement recommendation is proposed. The proposed method, which is based on machine learning methods like Decision Trees and Gradient Boosting along with the usage of Python and Pandas tries to predict crop yield based parameters such as type of crop, area under cultivation, Season, temperature, rainfall for the given field area in per hectare terms. Beyond the prediction, we make decision-support recommendations that could help boosting yield and profit. It also gives you best crop pairs, such as mirchi and onion which can be intercropped to get maximum benefit from the land and returns. The package also prescribes crop specific organic fertilizers and eco-friendly inputs that include neem oil and neem based formulations as the means for enhancing soil fertility, pest management while promoting sustainable agriculture farming. With predictive analytics and sound agriculturaladvice, the system provides all-around farmer-centred functionality, thereby improving productivity and prof- itability in modern farming.

***Index Terms*—Crop Yield Prediction, Machine Learning, De- cision Trees, Gradient Boosting, Agriculture, Profit Estimation**

I. INTRODUCTION

The food and economic security of particularly the devel- oping word, where farming is a principal occupation, depends largely on agriculture [10]. Reliable crop yield prediction supports farmers and policymakers in deciding what crops to plant, where, how much area did it cover and funding allocation and risk management respectively [7]. As the data become more available and powerful computational resources are also provided, machine learning has started to be used in agricultural issues [14]. These methods study historical and environmental data in order to identify patterns that impact agricultural output [10]. Therefore, crop yield estimation sys- tems are important tools that can contribute to increasing agricultural efficiency and mitigate unpredictability on the fields [11]. Most popular studies rely on machine learning techniques as Decision Trees, Random Forests and Gradient Boosting methods [11]–[13] to predict crop yield from climatic and agronomic parameters like rainfall,crop type, temperature, soil state,Season and cultivated area. Despite being highly accurate

tools to estimate yield, these models are restricted to achieve only high prediction performance [5]. Furthermore, there are few systems that provide prescriptive recommendations to farmers on how they can improve crop yield post-prediction [9]. This constraint limits the practical utility of such systems for farmers, which need not only forecasts but advice to im- prove productivity, profitability and sustainability in productive agriculture [8]. To address this limitation, this study introduces the development of an automatic crop yield prediction model with incorporation of predictive modeling and recommenda- tions for increased yielding. The model forecasts yield of crop depending on the type, area, Season, temperature, rainfall and district[11], [12]. The model also recommends ideal crop- pair combinations for intercropping (ie pairing mirchi with onion) to achieve most from the land [8]. The solutions also promotes crop- specific organic fertilizers, such as neem oil and neem-based formulations for sustaining soil health and farming system [9]. This harmonized action-oriented system becomes more farmer friendly and practical in use.

II. PROBLEM STATEMENT

The majority of current crop yield prediction methods concentrate on yield estimation rather than offering farmers advisory assistance to increase productivity [5]. The real- world applicability of current systems is limited because farmers are frequently given predicted yield values without advice on appropriate crop combinations or the use of organic fertilizer needed to increase yield [9]. Furthermore, due to a lack of knowledge and technological assistance, many farmers today are progressively disregarding conventional farming techniques like intercropping and organic fertilization [8]. This indicates a substantial research gap in the integration of practical agricultural advice with yield prediction. Therefore, there is a need for an intelligent decision-support system that forecasts crop yield while recommending suitable intercrop- ping strategies and organic fertiliZers to enhance productivity and promote sustainable agriculture [9].

III. OBJECTIVES

The overall goal of this project is to design an Automated Machine Learning system that can help accurately forecast crop yield based on agronomy, weather, and economy information, as well as provide prescriptions for how to increase crop productivity [11], [12]. The tool intends to help

the farmers by suggesting appropriate pair of crops for intercropping based on their growth periods as well as suggesting crop based organic fertilizers like neem-based formulations to add soil health and improve Sustainable farming practices [8], [9]. The project aims to enhance operational efficiency, economic viability and environmental sustainability of the system by combining yield prediction with decision-support advice, hence ensuring that real world farming applications can make practical use of it [14].

IV. LITERATURE REVIEW

Crop yield prediction has recently used machine learning more extensively, coping with the complexity of agricul- tural systems affected byclimatic, soil and agronomic effects[10],[11]. Earlier in 2021,the use of different machine learning models (e.g., regression, random forests and neural networks) was investigated to predict yield by including environmental factors such as rainfall, temperature and soil characteristics[10],[13]. Forexample, a research in 2021 used various ML models for the prediction of crop yields and displayed that structured dataset analysis serves as an important ground to increase forecast accuracy[10]. Researchers continued this foundational work from 2022 to 2024,with additional ensemble and hybrid models contributing to more robust and accurate prediction[11],[12],[14]. For dynamic yield prediction at regional scale and for multiple crops, time-series based methods using models including Random Forest, Gradient Boosting, XGBoost and bagging regressorsperformed best[3],[11]. Moreover, systematic reviews focused on recent progressin classical machine learning and deep learning methods for precision agriculture, which compared performance (e.g., R^2 RMSE) across different model architectures and crop types as well as challenges and limitations such as feature selection, data imbalance, and temporal dependencies[5],[14],[15].

V. RESEARCH GAP

The majority of current machine learning-based research on crop yield prediction solely concentrates on yield esti- mation using variables like crop type, rainfall, temperature, and area[10],[11],[12],[13],[15]. These systems primarily offer numerical forecasts without providing farmers with useful information[5],[14]. As a result, farmers are not given ad- vice on how to increase productivity, but they are informed about the anticipated yield. Yield prediction's practical utility in agricultural practices is limited because it is viewed as an outcome rather than a process that supports decision- making. Furthermore, recommendations for yield-improving techniques like crop combination (intercropping) and organic fertilizer selection are absent from current research[8],[9]. There is a big gap in the practical benefits of ML solutions for farmers due to the lack of integrated, farmer-centric advisory systems[5],[14].

VI. PROPOSED SYSTEM

The Automated Crop Yield Prediction System with Advi- sory Support that is being proposed is intended to help farmers make better decisions and increase yield [9]. This system offers three main functions, in contrast to conventional systems that only forecast crop yield [11]:

Fig. 1. Proposed system

- Crop yield prediction
- Intercropping recommendations
- Organic fertilizer recommendations

VII. METHODOLOGY

A. Dataset Collection

The data used in this research was obtained from publicly available agricultural data sources on Kaggle. The data holds information about past values of temperature, rainfall, type of crop, Season, District, area of cultivation[7],[10][15]. The data is in the form of structured CSV files and has been obtained from weather and agricultural sources. The choice of Kaggle for machine learning research was based on its reliability and ease of access[1],[2],[4].

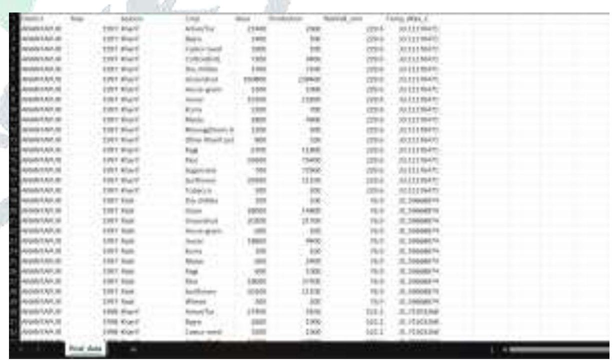

Fig. 2. Dataset for Crop Yield Prediction

B. Data Preprocessing

Data preprocessing uses Python and Pandas to enhance data quality and consistency [12]. Missing values are handled using the appropriate imputation techniques, and outliers and inconsistent records are removed. To preserve a consistent scale, numerical characteristics like temperature, rainfall, Area are normalised. Crop type,Season,District are categorical variables that are encoded into numerical formats that are appropriate for machine learning algorithms. In order to reduce model complexity and increase prediction accuracy, feature selection techniques are used to find pertinent attributes that significantly contribute to yield prediction [11].

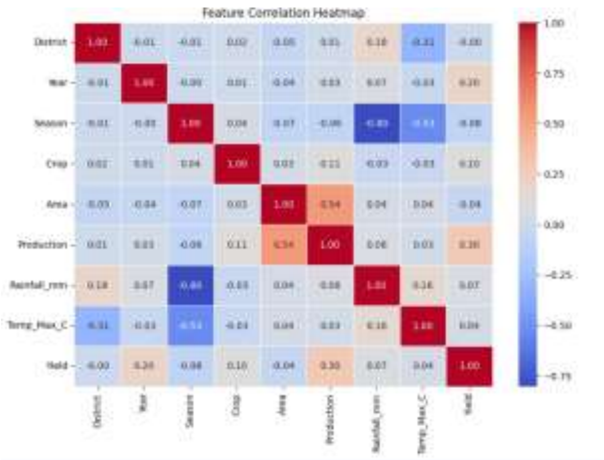

Fig. 3. Heatmap Helps identify which variables strongly affect Yield

C. System Overview

The system allows farmers to input parameters such as crop type, cultivated area, rainfall, temperature, Season, and District. Machine learning models analyze these inputs to predict yield while generating advisory recommendations [11], [12]. Forecast for Crop Yield Machine learning algorithms like Decision Tree and Gradient Boosting are used to predict crop yield [11], [12]. These models are implemented using Python and Pandas for feature handling and data preprocessing, and they are trained on historical agricultural data. 1. Type of crop 2. Rainfall on average 3. Season 4. area under cultivation 5. temperature After processing the input data, the trained model projects the expected yield, which the user can see.

Fig. 4. Dataset Used for Intercropping Recommendation

D. Crop Combination (Intercropping) Suggestions

The suggested system has an intercropping recommendation module to increase productivity [8]. The system recommends a compatible pair crop that can be grown in the same field when a farmer enters a particular crop type. To guarantee improved land use and higher yield, these crop combinations are chosen using agricultural expertise and compatibility guidelines [8]. For instance, it is advised to combine crops like mirchi with Onion to increase productivity.

E. Suggestion for Organic Fertilizer

Additionally, the system recommends organic fertiliser according to the crop that was chosen [9]. These suggestions support sustainable farming methods, lessen reliance on chemicals, and preserve soil fertility. The crop combination recommendations and fertilizer advice are shown in the system's advisory section.

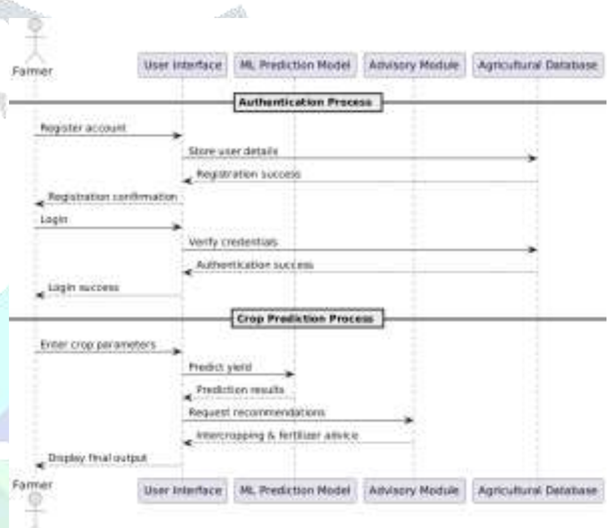

Fig. 5. Diagram Showing Interaction Between Farmer, ML Model, Database, and Advisory Module

F. Model Application and Prediction

Following preprocessing, crop yield predicted using ma- chine learning models like Decision Trees and Gradient Boost- ing [11], [12]. Decision trees capture non- linear relationships between input features and yield, whereas Gradient Boosting increases accuracy by combining multiple weak learners. The dataset is divided into training and testing sets in order to confirm model performance. To verify model performance, the dataset is split into training and testing sets. Prediction effectiveness is evaluated using metrics like mean squared error and R-squared score [11]. Based on inputs from the user, the trained models produce yield estimates. Crop Yield Prediction with Yield Enhancement suggestions 1: Register and login 2: Input crop and environmental parameters 3: Preprocess input data 4: Apply Decision Tree and Gradient Boosting 5: Predict yield 6: Generate advisory recommendations 7: Output results to farmer

G. *Advisory System*

In addition to forecasting, the system helps farmers in- crease productivity by offering advice [9]. The system rec- ommends appropriate intercropping combinations to improve land utilization based on the chosen crop and anticipated outcomes [8]. To enhance soil fertility and pest control, it also suggests eco- friendly inputs like neem oil and neem- based formulations as well as crop-specific organic fertilizers [9]. Predefined agricultural knowledge and historical trends are used to generate these recommendations. The system functions as a useful decision- support tool for farmers by combining prediction and advisory features.

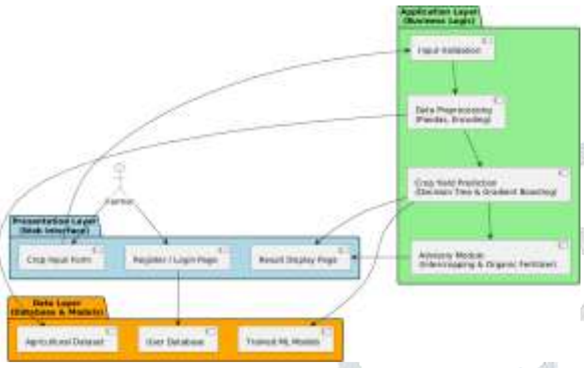

Fig. 6. Architecture Diagram of Automated Crop Yield Prediction and Advisory System

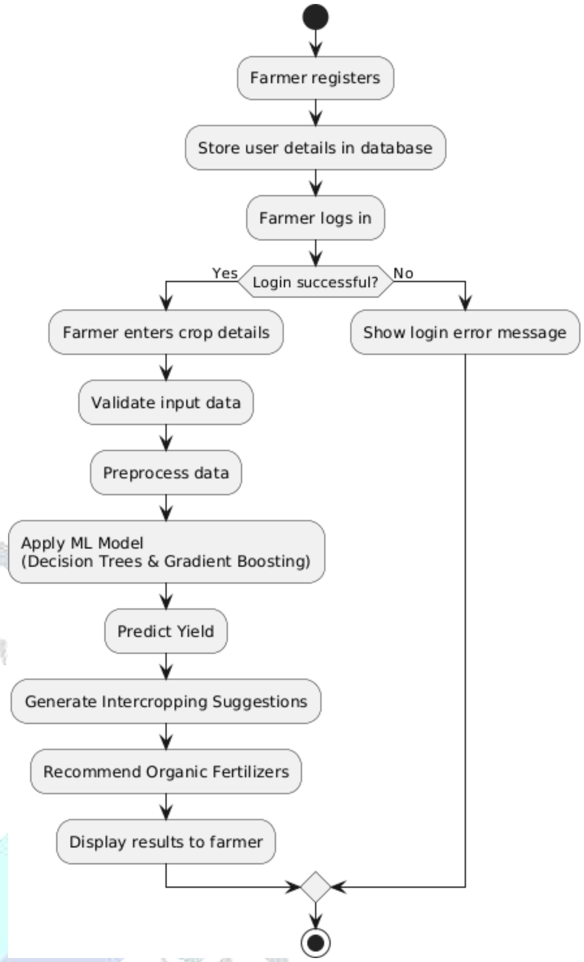

Fig. 7. Workflow of Crop Yield and Profit Prediction with Yield Enhancement Recommendations

VIII. RESULTS AND DISCUSSION

The proposed system was tested using past crop, envi- ronmental[1],[4],[7]. Machine learning algorithms like Deci- sion Trees and Gradient Boosting were employed to predict crop yield and the results were tested using R-squared (R^2) and mean squared error (MSE)[2],[11]. Gradient Boosting performed better than Decision Trees, with an R^2 value of 0.92 and lower MSE values, establishing high accuracy in predictions[11],[12]. Apart from prediction, the system also offers effective advisory support through recommendations of suitable intercrops like Onion with mirchi and organic fertilizers like neem-based formulations. All these aspects together increase productivity, profitability, and sustainability, making the system a complete farmer-supporting decision- making tool[5],[14].

IX. EXPECTED OUTCOMES

The creation of a machine learning-based system that can precisely forecast crop yield using variables like crop type, cultivated area, fertilizer use, temperature, rainfall, and Sea- son is the main anticipated result of this project [11]. The system is anticipated to provide accurate forecasts that assist farmers in estimating production levels and financial returns prior to cultivation by uing Decision Trees and Gradient Boosting algorithms [11], [12]. Better planning, effective resource allocation, and lower financial risk for farmers are all made possible by this predictive capability. In addition to forecasting, the initiative provides practical advisory support aimed at increasing crop productivity. To make better use of the land and increase yield, the system should recommend suitable intercropping combinations [8]. Apart from neem oil neem-based producer to promote soil fertility and sustainable agriculture practices also use of specific fertilizers like crop- based organic fertilizers are recommended [9]. It is expected that this system will reduce the gap between yield estimation and mitigation by integrating forecasting with practical advice to deliver a holistic, farmer-centric decision support service.

X. CONCLUSION

An automated crop yield prediction system that combines machine learning-based forecasting with farmer advisory sup- port was presented in this study. The system efficiently analy- ses agricultural and environmental factors, including crop type, cultivated area, Season, temperature, rainfall, and District, using Decision Tree and Gradient Boosting models [11], [12]. The experimental findings show that the suggested method offers accurate yield, allowing farmers to more effectively plan cultivation operations and manage resources [11]. The addition of recommend dation mechanisms that go beyond conventional yield prediction systems is a significant contribution of this work. In order to improve productivity and sustainability, the system recommends crop-specific organic fertilizers, such as neem-based formulations, and suggests appropriate crop pair combinations for intercropping [8], [9]. The suggested system closes the gap between analytical results and use- ful agricultural decision-making by fucusing prediction with practical advice. Enhanced farm profitability, productivity, and ecologically friendly farming methods are all supported by this integrate

REFERENCES

[1] K. Jaya Deepthi et al., "Smart Agriculture: Crop Recommendation and Yield Prediction Using Random Forest," Proc. ICITSM, 2025.

[2] V. Sawan et al., "Machine Learning Based Method for Forecasting Crop Yield," J. Recent Innovations in Computer Science and Technology, 2025.

[3] Y. Yan et al., "Crop Yield Time-Series Data Prediction Based on Hybrid Models," Preprint, 2025.

[4] S. Raju Sarikonda et al., "Agriculture Data Analysis and Crop Yield Prediction Using Machine Learning," Int. J. Adv. Eng. Hub, 2025.

[5] "Crop Yield Prediction Using Machine Learning: A Review," Smart Agricultural Technology, 2025.

[6] A. D. Kadam and N. S. Bansod, "Improving Crop Yield Prediction Using Machine Learning," J. Neonatal Surg., 2025.

[7] P. Patil et al., "Crop Selection and Yield Prediction Using ML," Curr. Agriculture Research J., 2023.

[8] "Intercropping Maize and Soybeans Improves Yield," BMC Plant Biol- ogy, 2024.

[9] C. Musanase et al., "ML-Based Crop and Fertilizer Recommendation System," Agriculture, 2023.

[10] S. Qiao et al., "Hierarchical Features for Crop Yield Prediction," IEEE
J. Sel. Topics Appl. Earth Observ. Remote Sens., 2021.

[11] R. Kumar and S. Verma, "Performance Analysis of Ensemble Learning Models for Crop Yield Prediction," *Procedia Computer Science*, vol. 218, pp. 162–169, 2022.

[12] A. Sharma and P. Kaur, "Machine Learning Approaches for Sustainable Agriculture and Crop Yield Prediction," *Journal of Cleaner Production*, vol. 330, pp. 129841, 2022.

[13] S. Qiao, J. Wang, X. Zhang, and Y. Zhang, "Crop Yield Prediction Using Multisource Data and Machine Learning," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 12345–12355, 2021.

[14] S. M. Shawon, F. B. Ema, A. K. Mahi, and H. T. Zubair, "Crop Yield Prediction Using Machine Learning: A Systematic Review," *Smart Agricultural Technology*, vol. 4, pp. 100182, 2022.

[15] M. Jeong, J. Kim, and S. Lee, "Crop Yield Prediction Using Weather and Soil Data with Machine Learning," *Computers and Electronics in Agriculture*, vol. 184, pp. 106093, 2021.