

AN ADVANCED PREDICTIVE ANALYTICS FRAMEWORK FOR CROP YIELD FORECASTING USING IOT AND MACHINE LEARNING

Mynashree B

PG,Student

Dept. of MCA

The Oxford College of Engineering,
Bommanahalli, Bengaluru- 560068
mynashreemca2025@gmail.com

Ashok B P

Assistant Professor

Dept. of MCA

The Oxford College of Engineering,
Bommanahalli, Bengaluru- 560068
ashokbp.mca@gmail.com

ABSTRACT

A key factor in maintaining food security and promoting economic growth is agricultural production. In order to meet difficulties presented by resource scarcity, climate change, and the expanding world population, accurate agricultural production forecasting has become more and more crucial. This research combines machine learning (ML) and the internet of things (IoT) to provide a sophisticated predictive analytics platform for crop yield forecasting. Real-time environmental and soil data, including temperature, humidity, rainfall, soil moisture, and nutrient levels, are gathered by the framework using IoT-enabled sensors. To guarantee quality and dependability, these data streams are sent to a centralized platform where preprocessing, feature engineering, and data fusion methods are used. Accurate yield estimates are produced by analyzing historical and current datasets using machine learning algorithms, such as regression models,

ensemble learning, and deep learning techniques.

KEYWORDS: Crop Yield Forecasting, Predictive Analytics, Internet of Things (IoT), Machine Learning, Smart Farming.

INTRODUCTION

Many economies rely heavily on agriculture for both food security and industrial raw materials. Crop production, however, is extremely susceptible to uncontrollable elements including water availability, pest infestations, soil fertility, and climate change. For farmers, legislators, and other stakeholders to make well-informed decisions about resource allocation, supply chain management, and food security planning, accurate crop production forecasting has become essential. Accuracy and responsiveness of traditional yield forecast techniques, which are based on manual observations and historical data, are frequently constrained by changing agricultural conditions.

Smart agriculture now has more options thanks to recent developments in machine learning (ML) and the internet of things (IoT). Real-time data on

temperature, humidity, rainfall, crop health, and soil moisture may be continuously collected by IoT sensors, creating a rich dataset for study. After processing this data, machine learning algorithms can find hidden trends, create predictive models, and more accurately anticipate crop yields. Farmers and agricultural specialists may predict yield outcomes, optimize farming techniques, increase productivity, and lower risks related to uncertainties by combining IoT-based data collection with sophisticated predictive analytics.

LITERATURE SURVEY

A major topic of study in field of agricultural analytics, crop yield forecasting aims to promote precision farming and increase food security. The nonlinear and dynamic elements influencing crop output were frequently missed by traditional forecasting methods, which mostly depended on statistical models and historical yield records. Recent research emphasizes the function of Internet of Things (IoT)-based smart farming systems, which allow for real-time sensor network monitoring of soil moisture, temperature, humidity, and nutrient levels. Predictive modeling requires large-scale, continuous data streams, which are provided by the integration of IoT devices with cloud platforms.

Concurrently, developments in machine learning techniques, including Random Forests,

Support Vector Machines, Gradient Boosting, and Deep Learning architectures, have demonstrated exceptional precision in simulating intricate connections between agricultural yields and environmental factors.

EXISTING WORK

Numerous studies have been conducted in field of crop production forecasting in recent years by combining sensor technologies and data-driven methodologies. Due to their incapacity to capture real-time environmental variables, statistical models and historical agricultural data have been mainstays of traditional yield estimation techniques, which frequently lack precision. Since the advent of the Internet of Things, scientists have concentrated on using wireless sensor networks to gather field-level data, including temperature, humidity, soil moisture, and crop health.

In parallel, large-scale datasets have been analyzed for yield prediction using machine learning approaches like regression models, support vector machines, and deep learning architectures.

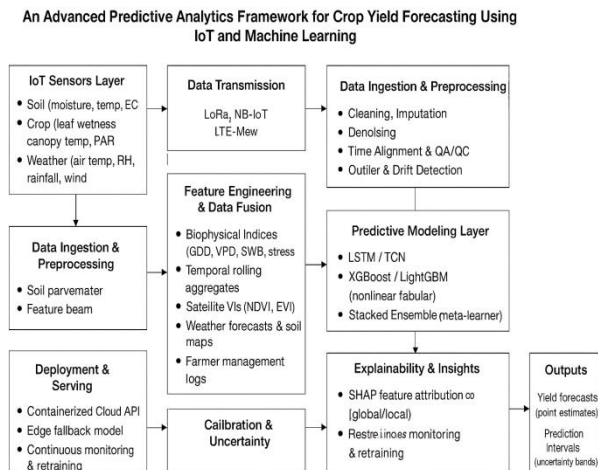
PROPOSED SYSTEM

The suggested method combines machine learning models with Internet of Things-enabled smart farming devices to present a sophisticated predictive analytics framework for agricultural yield predictions. Real-time data on important parameters like soil moisture, temperature, humidity, rainfall, pH level, and crop health indicators will be continuously collected by IoT sensors placed in agricultural areas. After being transferred to a cloud-based platform, this raw data will undergo preprocessing, feature extraction, and integration with previous agricultural datasets. Then, sophisticated machine learning techniques like Random Forest, Gradient Boosting, or Deep Neural Networks will be used to find correlations, evaluate trends, and produce precise crop yield forecasts.

In order to help farmers and other stakeholders make data-driven decisions about irrigation scheduling, fertilizer application, pest management, and harvesting, the system is made to deliver timely insights.

METHODOLOGY

By combining machine learning (ML) models with Internet of Things (IoT) devices, the proposed system forecasts crop yield with high accuracy and real-time capabilities. IoT sensors are initially deployed in agricultural fields to continuously collect environmental and soil data in real time, such as temperature, humidity, soil moisture, pH, and light intensity. This raw sensor data is sent to a cloud-based platform via wireless communication protocols (e.g., Wi-Fi, LoRa, Zigbee). Following storage, the data is preprocessed to ensure quality and consistency. This includes normalizing inputs, removing noise, and resolving missing values. Preprocessed dataset is then fed into supervised machine learning models like Random Forest or Support Vector Machine (SVM) or deep learning models like CNN or LSTM, depending on type of data. Environmental conditions and historical crop production data are used to teach these models patterns and relationships.



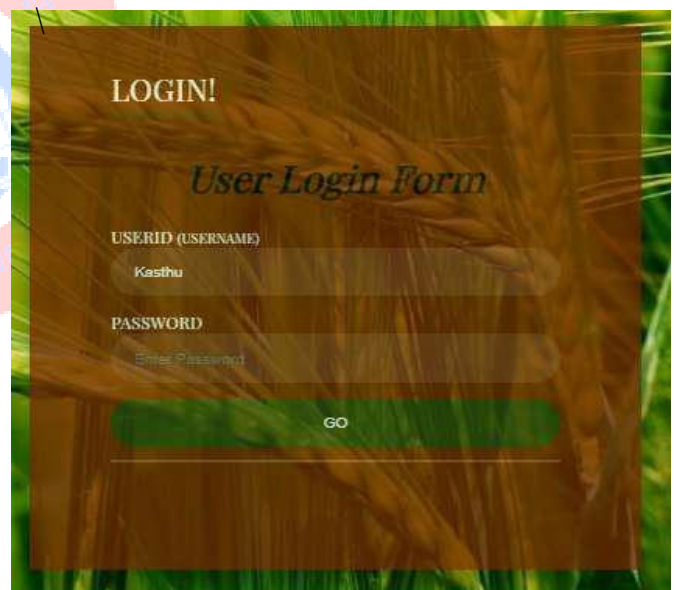
Model tuning and feature engineering are done to increase forecast accuracy. Trained model predicts expected crop production for a given time and place. System also features a feedback loop that considers new data from the field to continuously improve and update the model. It is optional to include a blockchain layer for data security and traceability. Result is a scalable, interpretable, and real-time yield prediction system that farmers can use via a mobile application or dashboard.

Hyperparameters are adjusted using cross-validation with grouped folds by field and season to avoid information leakage. System also incorporates quantile regression and calibration methods (Platt scaling, isotonic regression) to provide point forecasts and uncertainty intervals. Model is interpretable through the use of SHAP-based explainability, which enables the identification of significant drivers such as soil moisture or NDVI anomalies. Finally, trained model is deployed

via containerized APIs for cloud serving, and a lightweight surrogate model provides edge-level predictions in low-connectivity scenarios. With retraining triggered by seasonality or significant aberrations, system continuously monitors for data drift and performance degradation, ensuring long-term dependability and flexibility across crops and agroclimatic zones.

EXPERIMENTAL RESULTS

Through the End-User Login Page, registered farmers can safely enroll in the program. Enter your username and password to log in. Upon successful authentication, they are redirected to the dashboard.



- **The page for admin login**

System administrators can access and control the sophisticated predictive analytics framework for agricultural production forecasting through Admin Login page, which acts as a secure gateway.



Django administration

Username:

Password:

[Log in](#)

• **End User Signup/Register Page**

To access our cutting-edge predictive analytics platform, which is driven by real-time IoT sensor data and cutting-edge machine learning models, create a secure account. You will receive individualized insights by enrolling, including customized resource optimization techniques, soil moisture trends, and dynamic crop yield estimates.

Participation of ML in plantation in agriculture using decision making pattern

Welcome To Registration Page

Please fill in this form to create an account.

Clear All

Enter User ID :

Full Name

Contact

Security Question 1
 What was your childhood nickname?

Security Question 2
 What school did you attend for sixth grade?

Email

Password

Repeat Password

Gender Male Female Other

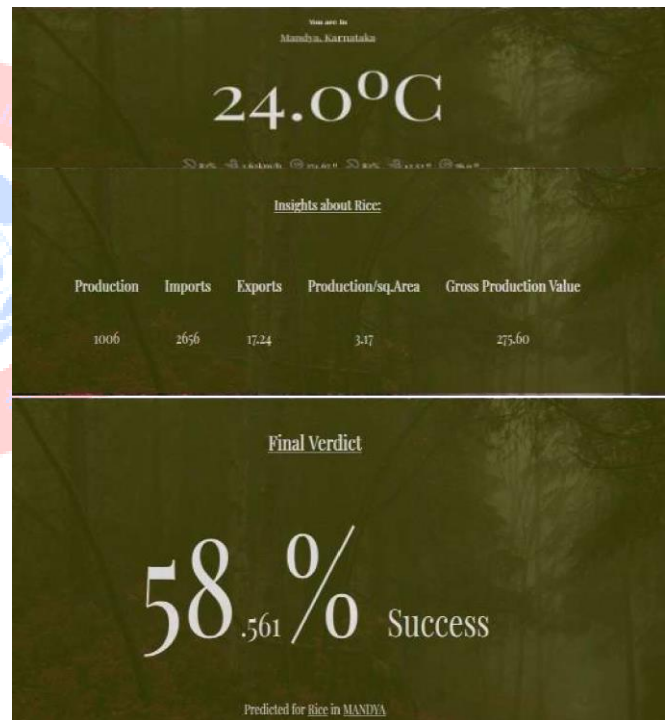
By creating an account you agree to our [Terms of Service](#)

[Sign Up](#)

[Already have an account? Sign In](#)

• **Result/Output Page for End User**

The end user's Result/Output Page is made to offer an intuitive, interactive, and clear interface that shows anticipated crop yield along with pertinent information. Accurate yield forecasts are produced by machine learning model once the IoT sensors have collected real-time agricultural data, including soil moisture, temperature, humidity, rainfall, and nutrient levels.



CONCLUSION

An intelligent and data-driven answer to problems facing agriculture industry is offered by the suggested advanced predictive analytics framework for crop production forecasting combining IoT and machine learning. Framework improves precision and dependability of yield estimation by using IoT-enabled sensors for real-time data collection and using machine learning algorithms for predictive modeling. By maximizing resource use, this technology not only gives farmers useful information to help them make better decisions about crop management, fertilizer use, and irrigation, but it also supports sustainable farming methods. Additionally, the framework's predictive power can help supply chain participants, agribusinesses, and policymakers guarantee food security and market stability. All things considered, the framework shows how combination of IoT and machine learning can transform conventional farming and open the door to future intelligent and sustainable farming.

REFERENCES

1. M. A. — "Machine learning for crop yield prediction: A comprehensive SLR," Heliyon (2024). ScienceDirect
2. "Crop yield prediction in agriculture: A thorough review of AI techniques," by S. Khan et al., Heliyon, 2024 (open access). PMC ScienceDirect
3. Frontiers in Plant Science, 2024; X. Liu et al., "An explainable Bi-LSTM model for winter wheat yield prediction." The frontiers
4. A. H. — "Computers and Electronics in Agriculture, 2022: Extreme Gradient Boosting for Yield Estimation Compared with Deep Learning."
5. USA crop yield estimation with MODIS NDVI," by J. Bolton et al., Remote Sensing, 2021. MDPI
6. In 2022, A. Zovko and colleagues published "Using Sentinel-2 data to estimate within-field wheat grain yield" in International Journal of Applied Earth Observation and Geoinformation.