iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** XI **Month of publication:** November 2025

DOI: https://doi.org/10.22214/ijraset.2025.75453

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

KhetAI: A Smart Farming Advisor

Kaustubh Khobragade[1] , Kartik Raut[2], Bhavesh Chawre[3], Gaurav Doye[4] , Dr. Ruchi Kadwane[5]
[1, 2, 3, 4]UG Student, Dept. Of Artificial Intelligence, GH. Raisoni University, Amravati, Maharashtra, India
[5]Assistant Professor. Dept. of Computer Science and Artificial Intelligence, GH. Raisoni University, Amravati, Maharashtra, India

Abstract: The system will integrate five machine learning models: Crop Recommendation, Crop Yield Prediction, Fertilizer Suggestion, Market Price Prediction, and Plant Disease Detection. Each of these uses supervised learning models such as Random Forest Classifier, Random Forest Regressor, XGBoost, and Convolutional Neural Networks in image-based disease diagnosis. The back-end is developed on FastAPI and Python while the React.js-based frontend allows farmers to input their native soil and environmental parameters in a user-friendly manner. Experimental results demonstrate high performance across models that have attained the accuracies of 95-99% in recommendation and prediction tasks.
Keywords: Agriculture, Artificial Intelligence, Crop Recommendation, Farming, Fertilizer Recommendation, Machine Learning, Market Price Prediction, Plant Disease Detection, Yield Prediction

I. INTRODUCTION

Farmers mostly depend on traditional knowledge or manual observation, which may not be accurate and/or sustainable in today's data-driven environment. KhetAI - A Smart Farming Advisor is an artificially intelligent decision support system designed to help farmers during the agriculture cycle. The overall architecture consists of five major models: Crop Recommendation, Crop Yield Prediction, Fertilizer Suggestion, Market Price Prediction, and Plant Disease Detection.

Together, these modules will help determine the answers to the following questions that are basic to agriculture: what to grow, how much, what fertilizer to use, what market price one should get, and how to detect plant diseases early. KhetAI makes use of machine learning algorithms. The models for predicting the crop and fertilizer use Random Forest Classifiers, the yield and market-price predictions make use of Random Forest and XG Boost Regressors, while the disease-detection component uses a Convolutional Neural Network (CNN) that has been trained on the Plant Village dataset. The system backend is implemented using Fast API and Python in model deployment and API integration, while the frontend is developed in React.js and tailwind to ensure a user-friendly web interface.

.

II. MATERIALS AND METHODOLOGY

A. System Architecture

The architecture is modular and data-centric in nature. These are the stages of the pipeline:

1) Data Collection: The dataset used in this work is obtained from authenticated sources such as Kaggle, Plant Village, and data.gov.in. The datasets have a wide variety of agricultural features including soil nutrients like Nitrogen, Phosphorus, and Potassium, temperature, humidity, pH, rainfall, along with market price statistics. In the disease detection part, labeled leaf image datasets have been used from Plant Village.
2) Data Preprocessing: Categorical variables are encoded with Label Encoding and One-Hot Encoding. After that, numerical variables are normalized using RobustScaler with the goal of reducing the impact caused by outliers. Outlier detection and correlation analysis are done to identify any redundant or noisy features.
3) Model Training and Evaluation: Each of the datasets is then divided into training and testing subsets, normally 75–80% for training and 20–25% for testing. For training different machine learning models, Python’s scikit-learn library is used, including Random Forest Classifier, Random Forest Regressor, and XGBoost RegressorFor the disease model, the CNN is implemented using the TensorFlow and Keras frameworks. Evaluation is done with Accuracy, R^2 Score, Mean Absolute Error, and Root Mean Square Error.
4) Integration and Deployment: All the models are trained and then serialized into .pkl (in the case of ML models) and .h5 (in case of CNN models) file formats, respectively. Further, these models are integrated into a backend created by using FastAPI that allows real-time predictions through RESTful APIs. The system is designed in a way to be extended for web and mobile platforms to make the insights more accessible and actionable by the farmers themselves.

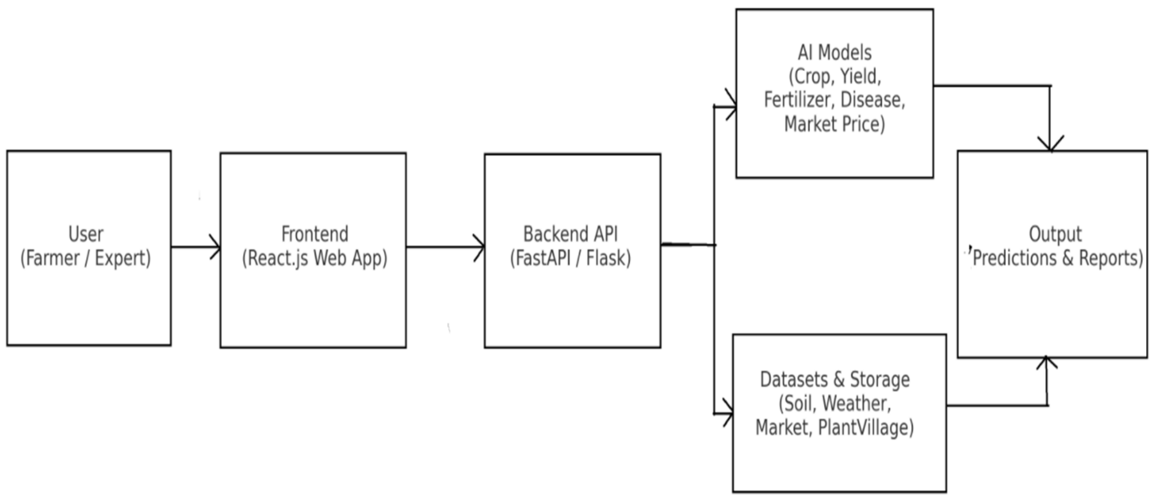


Fig.1: Integration workflow of KhetAI.

B. Crop Recommendation Module

Crop Recommendation is used for the best selection of crop based on their npk and many more parameters.

Dataset:

Crop_recommendation.csv it has 2200 samples with features that include N, P, K, temperature, humidity, pH, and rainfall .

Algorithm:

Random Forest Classifier

Methodology:

EDA is first performed on the given dataset by using a correlation heatmap and feature importance evaluation. Further, the data is divided into an 80:20 train-test split. The model is tuned for hyperparameters for better performance.

Result:

It achieved an accuracy of 98%. It works well in helping farmers decide which crop is most suitable for their soil with respect to weather conditions.

C. Crop Yield Prediction Module

The crop yield prediction system forecasts production based on environmental and historical factors.

Datasets:

Crop_yield.csv, predictioncrop.csv: Contains records of crop production, area, and climatic parameters across states and districts.

Algorithm:

Random Forest Regressor.

Methodology:

Yield = Production / Area, encoding categorical features State and Crop.

Tran and test split is 75:25.

Result:

Obtained $R^2 \approx 0.99$, therefore it is highly accurate prediction model that estimates the expected yield of crops.

D. Fertilizer Recommendation Module

Gives the best fertilizer recommendation based on soil npk paramters.

Dataset:

Fertilizer.csv contains Nitrogen (N), Phosphorus (P), and Potassium (K).

Algorithm:

Random Forest Classifier along with GridSearchCV for hyperparameter tuning.

Methodology:

The data standardization and then labels are encoded. The parameters, like the depth of trees and the number of estimators, are tuned using GridSearchCV.

Result:

Reached an accuracy of 98%. Hence, the system will correctly recommend the best fertilizer.

E. Crop Market Price Prediction Module

Dataset Source:

Primary data has been obtained from Government Open Data Platform, data.gov.in, and Agmarknet.

Algorithms Used:

Random Forest Regressor and XGBoost Regressor for price prediction.

Methodology:

Used GridSearchCV in 5-fold for hyperparameter tuning, such as tree depth, number of estimators, learning rate. Split the dataset into 80% training and 20% testing subsets. Ensured that all the crops and markets are represented and are balanced to avoid bias with respect to market trend.

Models evaluated using:

R^2: Coefficient of Determination, to measure overall fit.

RMSE: root mean square error to check the accuracy of the prediction.

High R^2, low RMSE and MAE values confirmed the strong predictive performances.

F. Plant Disease Detection Module

This module automatically detects crop diseases by classifying image of the leaf.

Dataset:

PlantVillage Dataset contains thousand of image of healthy and diseased leaves for various kinds of crop.

Framework:

TensorFlow and Keras to build a CNN architecture.

Architecture Details:

Three convolutional blocks comprising Conv2D, Batch Normalization, MaxPooling, and Dropout layers. Dense, SoftMax output layer. The size of the input is standardized to 128×128×3 pixels.

Training Parameters:

Optimizer: Adam.

Loss Function:

for this Sparse Categorical Cross entropy.

Metrics:

for metrics Accuracy, Top-3 Accuracy

Data Augmentation:

for proper understanding of data, Rotation, flipping, zoom, and brightness adjustments is used.

Result:

Achieved 49% training accuracy and 43% validation accuracy in four epochs. The model will perform even better with more epochs and using transfer learning.

G. Integration Methodology

Integration in the case of KhetAI ensures the smooth interaction between the now-trained machine learning models with the backend API layer and the user-facing frontend interface.

Backend Integration:

The trained models will be integrated into the backend using the Flask and FastAPI framework, in which all predictive services would be mapped to a dedicated RESTful API endpoint. The serialized models .h5 and .pkl files will be loaded into memory at backend initialization so that instant inferences can be enabled. Each endpoint, receives input in structured JSON format. Backend is hosted over Render.

Frontend Communication:

These APIs also interact with the frontend, which is developed using React and Tailwind CSS. Frontend is hosted over Vercel. CORS is enabled to ensure secure communication between the backend hosted on Render and the frontend hosted on Vercel.

III. RESULTS AND DISCUSSIONS

The KhetAI Smart Farming Advisor is an integrated platform of several AI-based predictive and classification models that come together to form a holistic decision-support system for farmers. Every module is trained, tested, and validated individually with its respective dataset. The results showed that the system performs with high accuracy, stability, and potential in real-world agricultural applications. What follows is a summary and discussion of the results derived from each model.

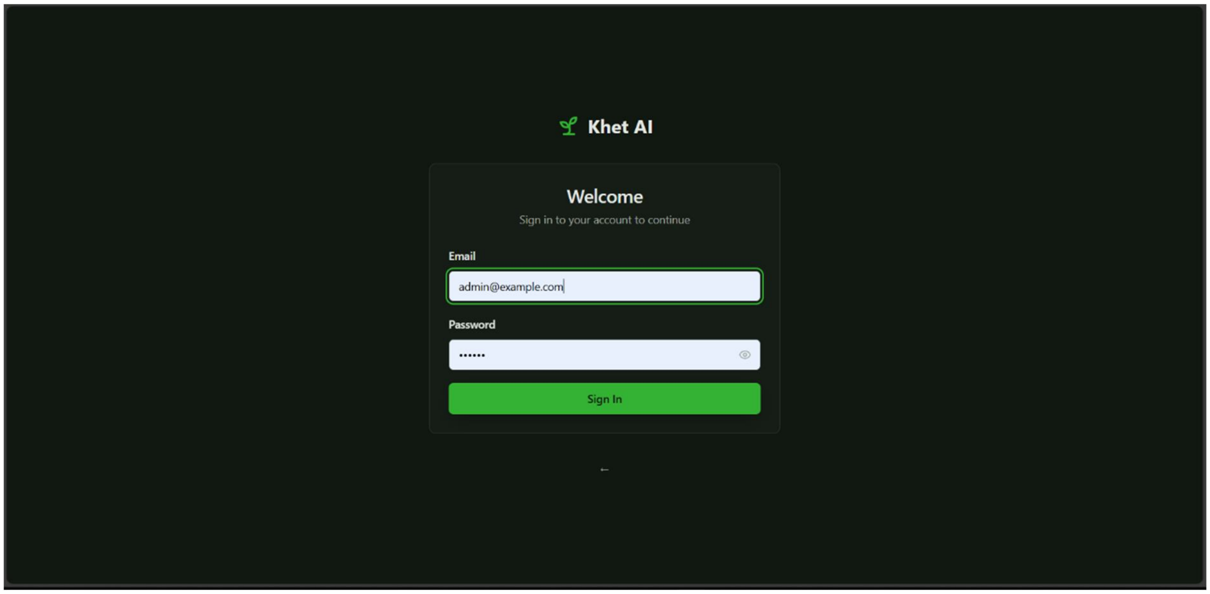


Fig.2: Login page of KhetAI

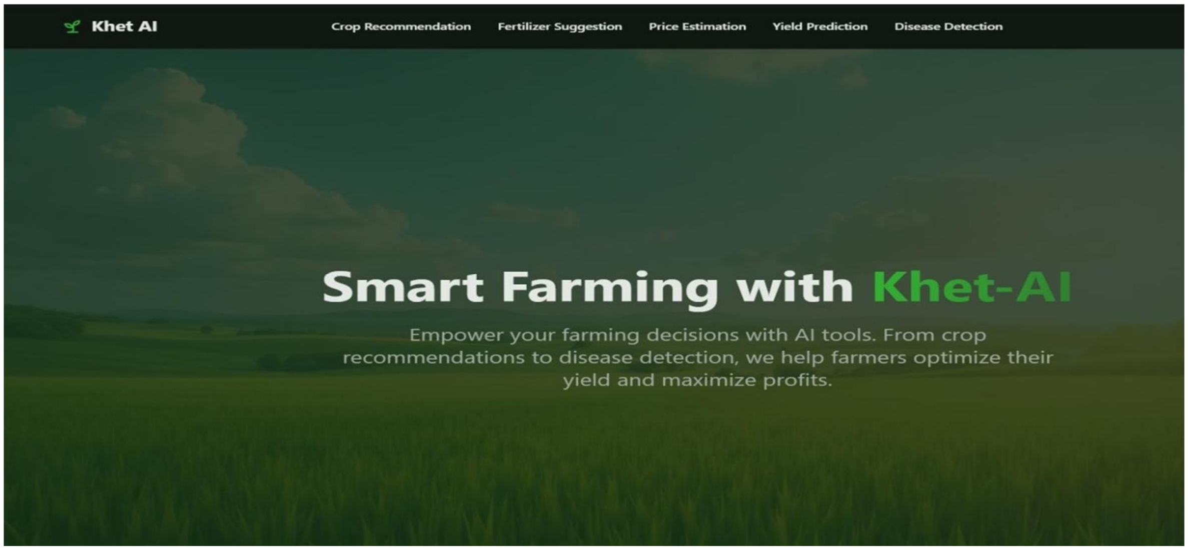


Fig.3: KhetAI Dashboard.

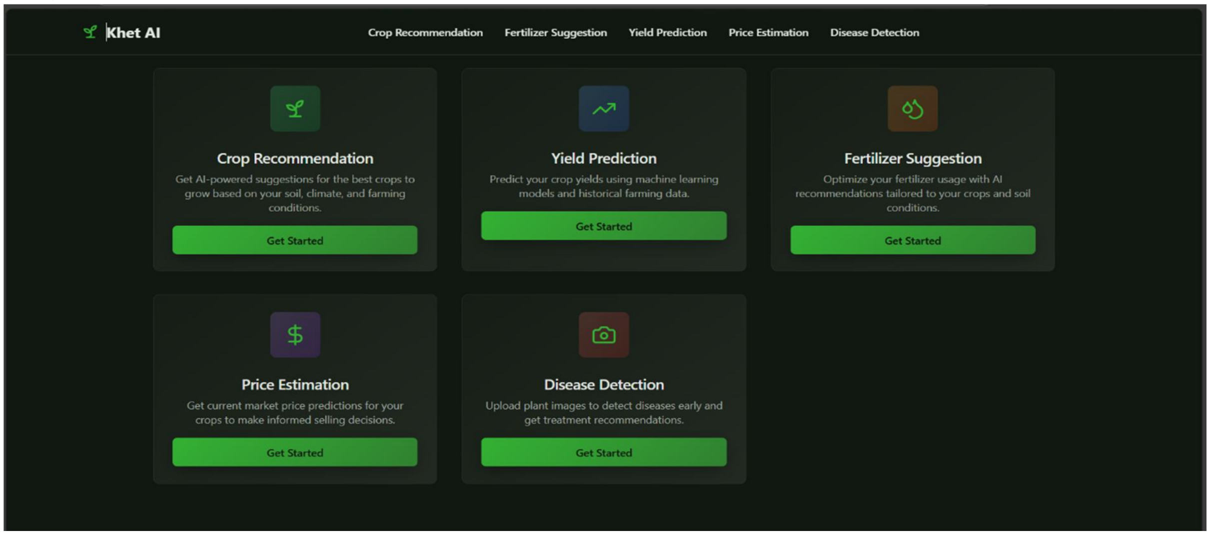


Fig.4: KhetAI dashboard (multiple paths to choose).

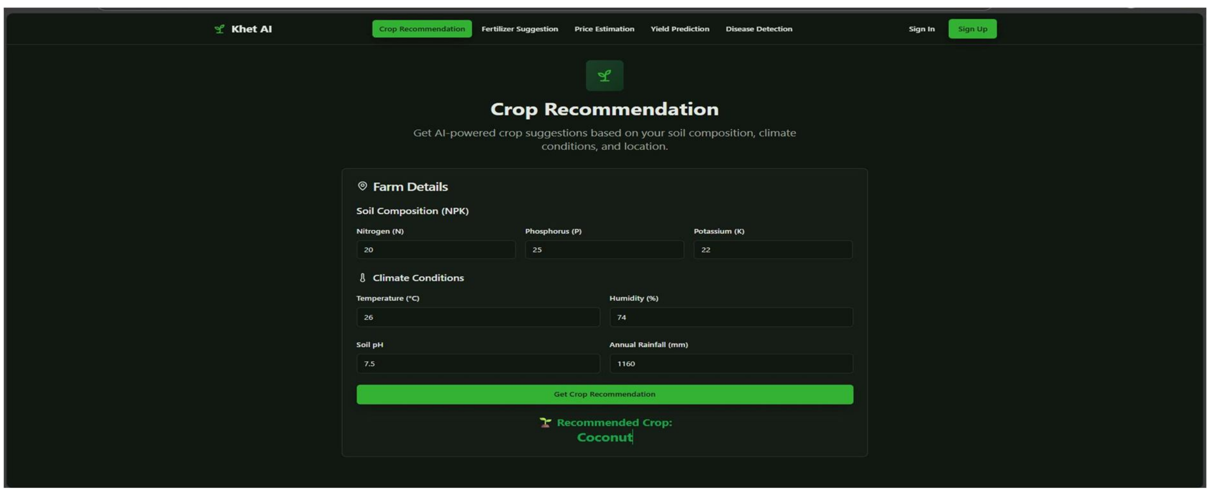


Fig.5: Crop recommendation

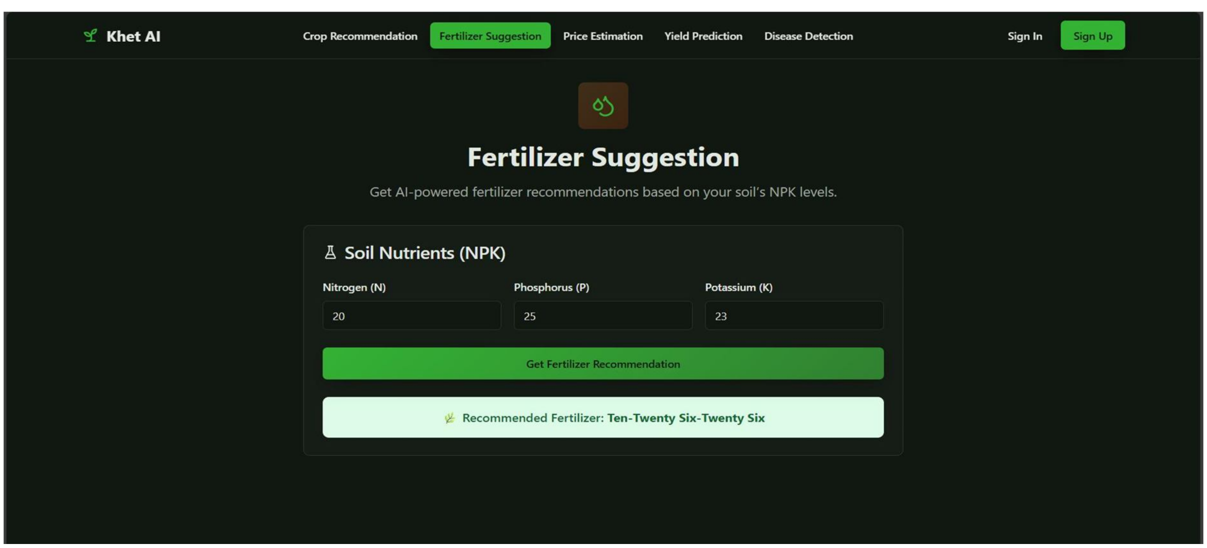


Fig.6: Fertilizer recommendation.

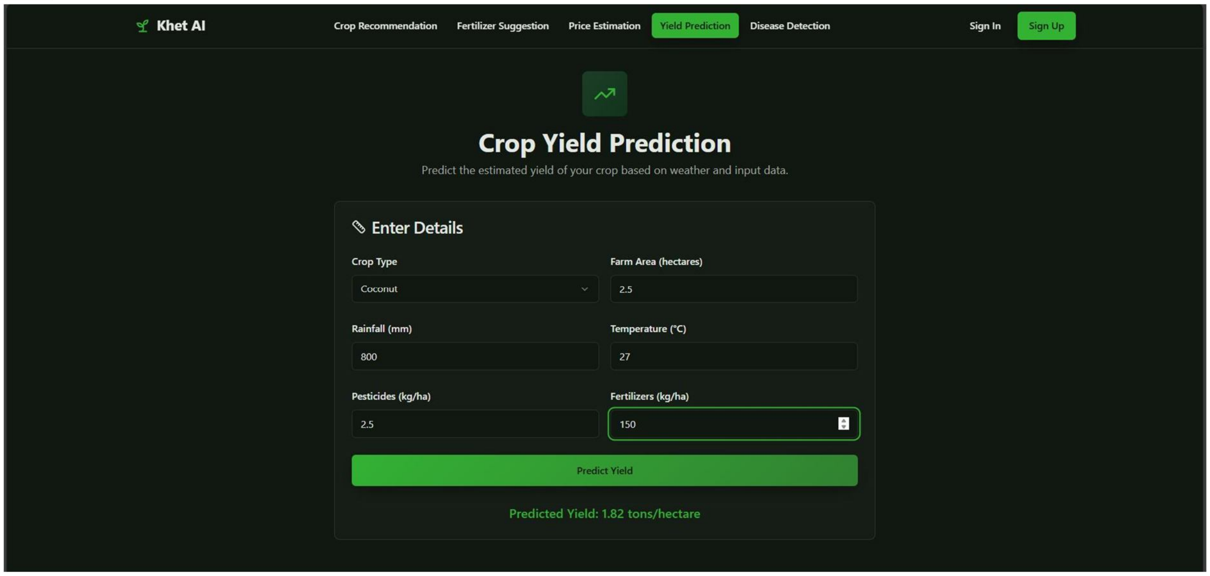


Fig.7: Crop yield prediction.

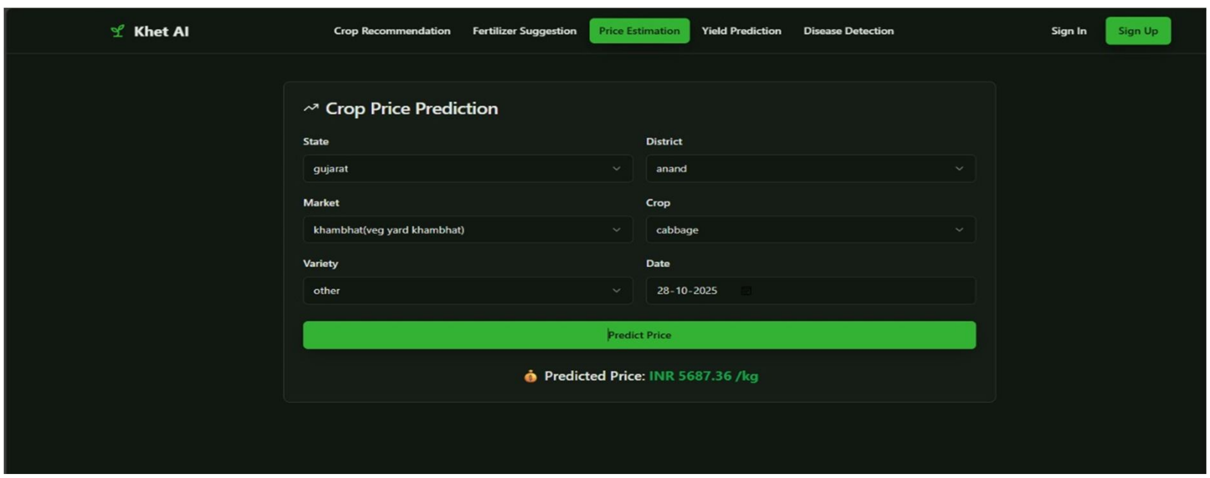


Fig.8: Crop price prediction.

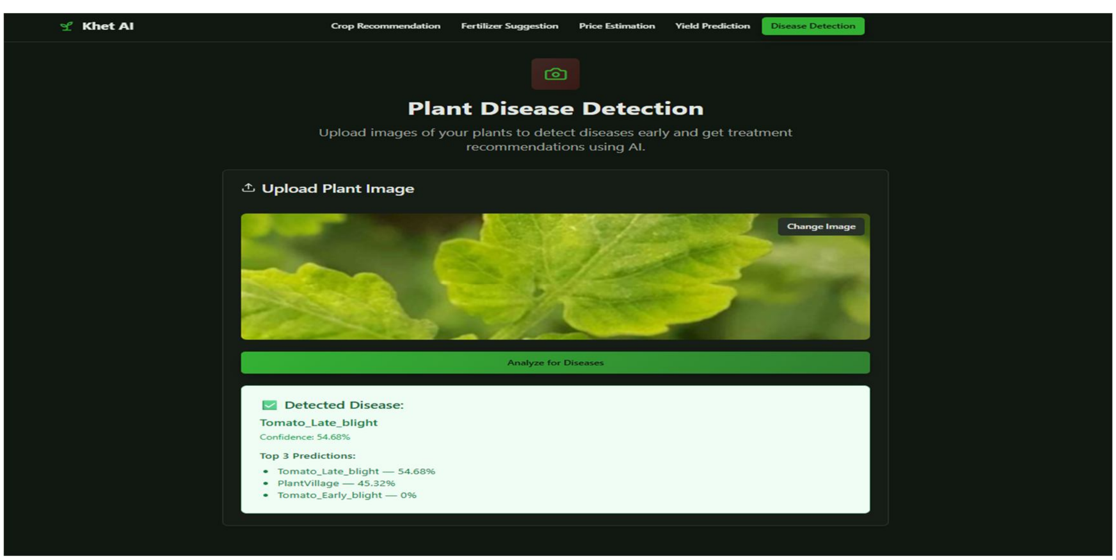


Fig.9: Plant disease detection.

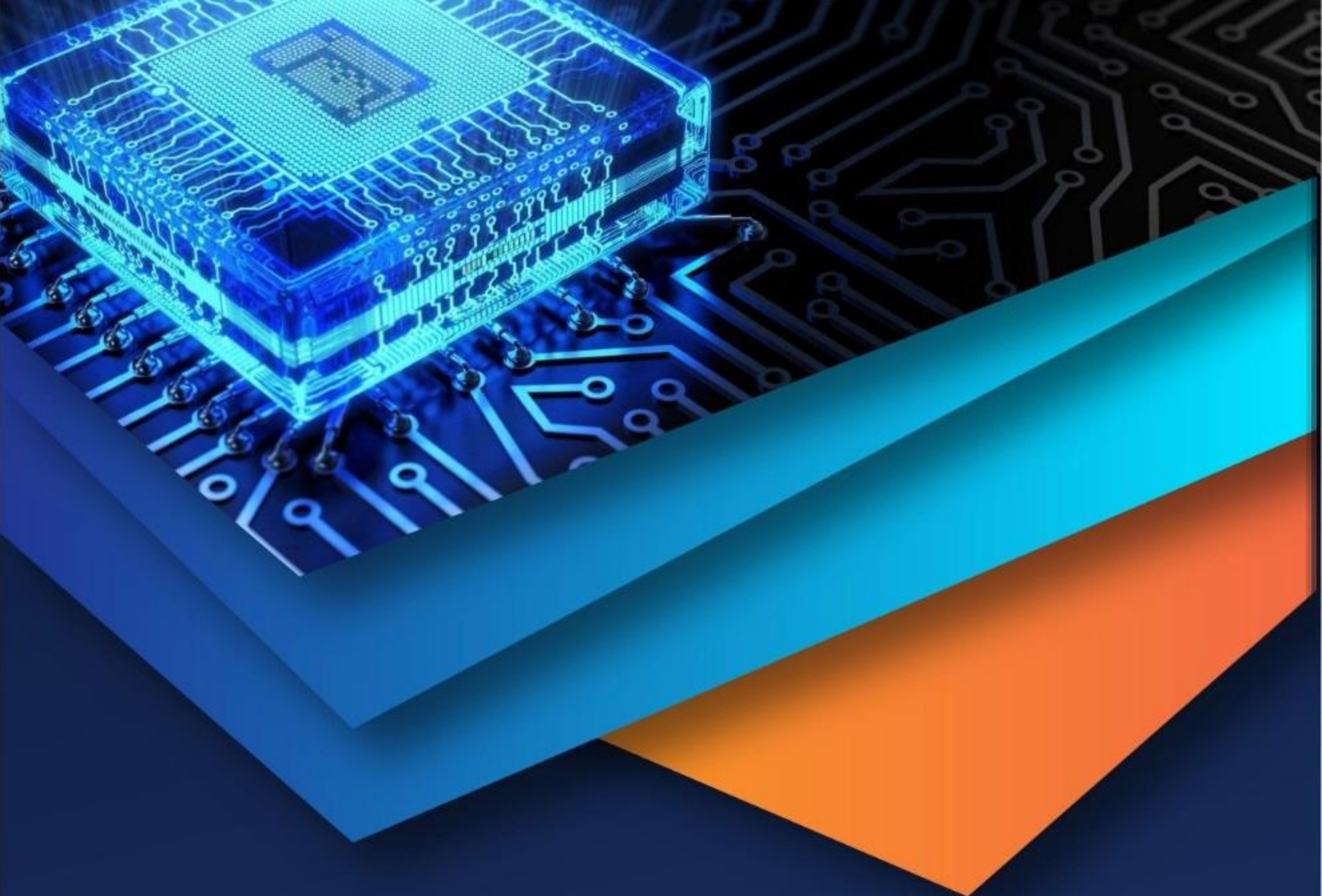
International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

IV. CONCLUSION

The KhetAI system provides an integrated intelligent platform for artificial intelligence and machine learning in acquiring data-driven insights in agriculture. Integrating these predictive models on Crop Recommendation, Yield Prediction, Fertilizer Suggestion, Market Price Forecasting, and Plant Disease Detection into one web-based decision-support system, the KhetAI effectively addresses main challenges faced by farmers in crop planning, disease management, and market decisions.

REFERENCES

[1] J. Singh, A. Sharma, and P. Patel, “Machine Learning-Based Crop Recommendation System for Precision Agriculture,” International Journal of Advanced Computer Science, vol. 12, no. 4, pp. 221–230, Apr. 2022.
[2] PlantVillage Dataset – Open Source Leaf Disease Image Dataset, Pennsylvania State University. Accessed: Jan. 2025.
[3] Kaggle, “Crop Recommendation Dataset,” Kaggle Open Data Platform. Accessed: Jan. 2025.
[4] Government of India, “Agmarknet – Agricultural Produce Market Price Data,” Open Government Data (OGD) Platform India. Accessed: Jan. 2025.
[5] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), San Francisco, CA, USA, 2016, pp. 785–794.
[6] F. Chollet, “Keras: The Python Deep Learning API,” GitHub Repository, 2015. Accessed: Jan. 2025.
[7] FastAPI Documentation, “FastAPI: High performance web framework for building APIs with Python.” Accessed: Jan. 2025.
[8] Meta Platforms Inc., “ReactJS Documentation.” Accessed: Jan. 2025.
[9] S. R. Chaurasia, “Crop Yield Prediction Using Random Forest Regression,” in IEEE International Conference on Computational Techniques, 2023, pp. 507–512.
[10] R. Hasan, “AI-driven Market Price Forecasting System Using Regression Models,” Journal of Agriculture Informatics, vol. 15, no. 3, pp.



10.22214/IJRASET

INDEX COPERNICUS

45.98

ISRA JIF

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY