

AI for Smart Agriculture using Machine Learning

Yogeshwari Deshmukh, Dr. Anita Mahajan

U.G. Student, Department of Computer Engineering, Ajinkya DY Patil school of Engineering, Pune,
Maharashtra, India

Associate Professor, Department of Computer Engineering, Ajinkya DY Patil School of Engineering, Pune,
Maharashtra, India

ABSTRACT: Agriculture plays a vital role in India's economy and food security but faces major challenges such as climate change, soil erosion, and inefficient resource use. To address these issues and enhance productivity, accurate crop yield prediction and fertilizer recommendations are essential. This study uses a sensor-based dataset from Kaggle containing key parameters like NPK levels, pH, temperature, humidity, rainfall, soil type, and crop type. Machine learning models such as Linear Regression (LR) and Artificial Neural Networks (ANN) were applied to predict crop yields, while a fuzzy logic approach recommended suitable fertilizers based on soil and climate conditions. This hybrid AI-driven system improves yield forecasting accuracy and promotes sustainable farming by reducing fertilizer overuse.

KEYWORDS: Machine Learning, NPK, Fertilizer, Crop yields.

I. INTRODUCTION

Agriculture has long been the backbone of economic stability and food security in developing nations like India. However, traditional farming methods often rely on experience and intuition rather than data-driven insights. With the emergence of Artificial Intelligence (AI), Machine Learning (ML), and Internet of Things (IoT) technologies, agriculture is undergoing a digital transformation. These innovations enable farmers to make informed decisions about crop selection, irrigation, and fertilizer use based on real-time soil and environmental data, leading to improved productivity and sustainability.

II. LITERATURE SURVEY

In [1] presents a machine learning-based model for accurately forecasting crop yield by analysing environmental and soil factors. The authors utilize regression techniques to process large datasets and identify patterns that influence production rates. Their system enhances agricultural decision-making by providing data-driven insights to farmers, thereby improving productivity and sustainability.

According to [2] the introduces a machine learning framework designed to recommend suitable crops for cultivation based on soil nutrients, weather, and regional conditions. By employing algorithms like Decision Trees and Random Forests, the model predicts the most profitable and sustainable crop choices. The study demonstrates improved accuracy and reliability in agricultural planning compared to traditional methods.

According to [3] the integrates yield and price forecasting into a single application to aid farmers in market-oriented decision-making. Using supervised learning algorithms, the system analyzes historical production and market data to predict both yield quantity and expected market prices. The study emphasizes the potential of data analytics in stabilizing farmer incomes and optimizing resource use.

According to [4] explores recent advancements in machine learning applications for crop yield prediction. It reviews diverse algorithms such as neural networks, support vector machines, and ensemble models used in agriculture. The authors highlight key environmental, genetic, and climatic factors influencing yield, offering a comprehensive view of current methodologies and future research challenges.

According to [5] the authors propose a comparative analysis of various machine learning algorithms—such as Linear Regression, KNN, and Random Forest—for estimating crop yields. Their experiments on real agricultural datasets

reveal that ensemble methods outperform conventional approaches in prediction accuracy. The study suggests that intelligent modeling can enhance farming efficiency and production forecasting.

According to [6] an implements supervised learning algorithms to estimate crop yields based on climatic and soil parameters. The model processes past yield data to establish relationships among input variables affecting productivity. Results show that the machine learning approach significantly improves prediction precision, aiding better crop management strategies.

According to [7] the proposes a hybrid predictive model combining machine learning and deep learning techniques. By integrating methods such as Random Forest and Convolutional Neural Networks (CNN), the framework enhances the precision of yield forecasts. The hybrid model captures complex non-linear dependencies in agricultural datasets, outperforming single-algorithm systems.

According to [8] the research examines how machine learning can be applied to predict historical and future crop yields under different climate scenarios. It leverages diverse datasets to train models that assess temporal variations in productivity. The study underscores the potential of ML in improving agricultural forecasting amid changing environmental conditions.

According to [9] evaluates multiple machine learning models across various datasets to determine their efficiency in yield prediction. The authors explore the scalability and generalization ability of algorithms when applied to different crops and regions. Their results indicate that advanced ensemble models consistently deliver higher predictive performance and adaptability.

According to [10]T an IoT-based system for monitoring soil nutrients in real time to support precision agriculture. The framework collects and analyses soil data using sensors, enabling the model to recommend the most suitable crops for specific soil conditions. By integrating machine learning algorithms, the system improves accuracy in crop selection and optimizes resource utilization. The approach demonstrates potential to enhance productivity, reduce wastage, and support sustainable farming practices.

Table 1: Comparison table for the literature review

Sr. No	Methodology	Algorithm Used	Limitations
11	ML-based crop recommendation considering NPK, soil pH, and climatic variables in India	Decision Trees, Random Forest, SVM	Limited regional focus; may not generalize to other climates or soil types
12	Fertilizer recommendation using ML models	Gradient Boosting, Neural Networks, Linear Regression	Dependent on quality of soil data; may require frequent model retraining for accuracy
13	NPK soil fertility measurement and crop recommendation system	KNN, Random Forest	Focuses mainly on soil nutrients; ignores other factors like pest or disease prevalence
14	Smart farming integrating IoT and ML to enhance crop yield	Random Forest, ANN, Regression Models	High hardware cost; sensor calibration and network reliability can affect performance
15	Sustainable crop recommendation using NPK sensors	Decision Trees, Support Vector Machines	Limited dataset size; may not account for long-term soil fertility changes

A. Gap Analysis

- Most studies focus on historical data and specific crops/regions; limited adaptability to real-time environmental changes or market dynamics [1], [3], [5], [6], [9].
- While soil nutrient-based recommendations are effective, they often do not integrate climatic variations or IoT enabled real-time data fully [2], [10], [11], [13], [15].

- Hybrid models improve prediction accuracy but require high computational resources and extensive datasets for training [7], [8].
- Integration of IoT sensors with ML enhances real-time decision-making, yet challenges exist in sensor calibration, network reliability, and cost-effectiveness[10] ,[14]
- Deep learning models capture complex non-linear relationships but may overfit small datasets and lack interpretability for end-users [4] [15]
- Model accuracy depends on high-quality soil data and frequent updates; may not generalize across regions with varying soil types [12]

III. CHALLENGES IN DATA

- Farmers may be reluctant to share farm-specific data due to privacy or economic concerns.
- Limited access to data can hinder large-scale adoption and model validation.
- Soil fertility, weather, and pest/disease conditions change over time, making static datasets less useful for accurate predictions.
- Continuous updating of datasets is necessary, but this increases complexity in storage and model retraining.
- IoT-enabled precision agriculture requires real-time monitoring of soil nutrients, moisture, and climate variables.
- Network latency, sensor calibration, and power constraints can limit continuous and reliable data collection.
- Data comes from multiple sources, such as IoT sensors, satellite imagery, weather stations, and market reports.
- Integrating these heterogeneous datasets into a unified format for machine learning is challenging and can lead to inconsistencies.
- Agricultural datasets often contain missing, inconsistent, or noisy values due to sensor errors or manual data entry.
- Low-quality data can significantly reduce the accuracy of machine learning models for yield prediction and crop recommendation

IV. APPLICATIONS OF MACHINE LEARNING TECHNIQUES ON DIFFERENT DATASETS

Machine learning techniques have been widely applied to diverse agricultural datasets to enhance crop yield prediction, recommendation systems, and precision farming. Historical crop yield data is commonly analyzed using algorithms such as Random Forest, Support Vector Machines, and Gradient Boosting to forecast production and optimize resource allocation.

Table 3: Overview of Machine and Deep Learning Applications on Different Datasets

Sr. No	Dataset Description	ML & DL Technique / Algorithm Used
1	Crop yield dataset including environmental and soil parameters	Regression models, Random Forest, SVM
2	Soil and climatic data for crop recommendation	Decision Trees, Random Forest, KNN
3	Historical crop yield and market price data	Linear Regression, Random Forest, Gradient Boosting
4	Survey of ML approaches for crop yield prediction	Review of ML algorithms including Neural Networks, Ensemble Methods
5	Multi-crop agricultural yield datasets	Linear Regression, Random Forest, KNN
6	Historical yield and environmental datasets	Regression models, ANN, Random Forest
7	Crop yield datasets integrating multiple features	Hybrid ML + Deep Learning (Random Forest + CNN/ANN)
8	Past and future crop yield data with climatic features	Neural Networks, Random Forest, Gradient Boosting

9	Diverse crop yield datasets across regions	Ensemble ML models, Random Forest, SVM
10	IoT soil nutrient sensor data	Decision Trees, Random Forest, ML algorithms
11	Soil NPK, pH, and climate data for Indian crops	Random Forest, Decision Trees, SVM
12	Fertilizer application and soil dataset	Gradient Boosting, Neural Networks, Regression Models
13	NPK soil fertility measurements	KNN, Random Forest
14	IoT-enabled smart farming datasets	ANN, Regression Models, Random Forest
15	Soil NPK sensor dataset for sustainable crop recommendation	SVM, Decision Trees, Neural Networks

V. DISCUSSION: TRENDS

- ML algorithms like Random Forest, SVM, and Gradient Boosting are widely applied for crop yield prediction and recommendation.
- Hybrid models combining ML and deep learning are emerging to capture complex non-linear relationships in agricultural data.
- IoT-enabled monitoring of soil nutrients, moisture, and environmental factors is becoming popular for real-time crop management.
- Many systems focus on NPK and soil pH analysis for recommending suitable crops and fertilizer application.
- Integration of historical yield, price, and climatic data is supporting smarter farming decisions and resource optimization.
- There is a need for models that adapt to changing environmental and climatic conditions for improved prediction accuracy.
- Increasing emphasis on sustainable practices, including efficient fertilizer use and crop rotation strategies, guided by ML insights

VI. FUTURE RESEARCH DIRECTIONS

- Combine IoT sensors, satellite imagery, weather data, and market trends to improve model accuracy and reliability.
- Explore hybrid and ensemble models to capture complex patterns in crop yield and recommendation systems.
- Develop models that update dynamically with changing soil, climate, and pest/disease conditions for timely decision-making.
- Incorporate price and demand forecasting alongside yield predictions to help farmers make economic decisions.
- Use cloud computing to manage large agricultural datasets and facilitate real-time processing and accessibility.
- Enhance sensor accuracy, calibration, and network reliability for continuous precision agriculture monitoring.
- Focus on environmentally friendly farming practices, including efficient fertilizer use, crop rotation, and water management guided by ML models.
- Design models that can generalize across different crops, regions, and climatic conditions, increasing practical applicability.

VII. CONCLUSION

The implementation of machine learning (ML) in agriculture for crop yield prediction and crop recommendation presents a significant leap forward in optimizing farming practices, improving yields, and promoting sustainable agriculture. By utilizing historical data inputs from agricultural datasets and machine learning algorithms, farmers can make informed decisions that reduce costs, increase productivity, and mitigate risks posed by environmental uncertainties.

REFERENCES

- [1] Mr. Telise Vinod,N. Chandrasena Reddy,Rapolu Suresh Reddy,Anugu Athrey Reddy. "Crop Yield Prediction using Machine Learning" International Journal of Engineering Research & Technology (IJERT) IJERTV12IS040077 Vol. 12 Issue 04, April-2023.
- [2] Shafiulla Shariff,Shwetha R B,Ramya O G,Pushpa H,Pooja K . "Crop Recommendation using Machine Learning Techniques" International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 ICEI - 2022 Conference Proceedings Special Issue - 2022
- [3] Prameya R Hegde, Ashok Kumar A R . "Crop Yield and Price Prediction System for Agriculture Application" International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 IJERTV11IS070060 Vol. 11 Issue 07, July-2022
- [4] Bali, Nishu, and Anshu Singla. "Emerging trends in machine learning to predict crop yield and study its influential factors: A survey." Archives of computational methods in engineering (2022): 1-18.
- [5] Gunjan, Vinit Kumar, et al. "Prediction of agriculture yields using machine learning algorithms." Pro- ceedings of the 2nd International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications: ICMISC 2021. Springer Singapore, 2022.
- [6] Reddy, D. Jayanarayana, and M. Rudra Kumar. "Crop yield prediction using machine learning algo- rithm." 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2021.
- [7] Agarwal, Sonal, and Sandhya Tarar. "A hybrid approach for crop yield prediction using machine learning and deep learning algorithms." Journal of Physics: Conference Series. Vol. 1714. No. 1. IOP Publishing, 2021.
- [8] Morales, Alejandro, and Francisco J. Villalobos. "Using machine learning for crop yield prediction in the past or the future." Frontiers in Plant Science 14 (2023): 1128388.
- [9] Joshua, S. Vinson, et al. "Crop yield prediction using machine learning approaches on a wide spectrum." Computers, Materials & Continua 72.3 (2022).
- [10] Senapaty, Murali Krishna, Abhishek Ray, and Neelamadhab Padhy. "IoT-enabled soil nutrient analysis and crop recommendation model for precision agriculture." Computers 12.3 (2023): 61.
- [11] Dey, Biplob, Jannatul Ferdous, and Romel Ahmed. "Machine learning based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables." Heliyon 10.3 (2024).
- [12] Tanaka, Takashi ST, et al. "Can machine learning models provide accurate fertilizer recommendations?." Precision Agriculture 25.4 (2024): 1839-1856.
- [13] Okwii, Simon. "NPK soil fertility measurement system and crop recommendation based on machine learning." (2023).
- [14] Sundaresan, S., et al. "Machine learning and IoT-based smart farming for enhancing the crop yield." Journal of Physics: Conference Series. Vol. 2466. No. 1. IOP Publishing, 2023.
- [15] Gottemukkala, Lavanya, et al. "Sustainable crop recommendation system using soil npk sensor." E3S web of conferences. Vol. 430. EDP Sciences, 2023.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details