"Smart Agriculture through Machine Learning" Deployable Model for Multi-Aspect Farm Advisory

Mr.P.Bharath Krishna[1], Dr K Gajalakshmi[2]
[1.] *Senior Manager, HPE, Bangalore,* [2.] *Assistant Professor, RV College of Engineering, Bangalore*

***Abstract*—Smart agriculture is increasingly essential in contemporary farming due to challenges such as climate fluctuations, volatile market conditions, and suboptimal use of agricultural resources, all of which negatively impact productivity. This work introduces a practical and implementable machine learning–driven framework titled "Smart Agriculture through Machine Learning: A Multi-Aspect Farm Advisory System." The proposed system consolidates four key functionalities—crop recommendation, yield estimation, market price prediction, and fertilizer advisory—within a single decision-support platform. Each component employs data-centric machine learning techniques trained on soil properties, weather patterns, historical pricing data, and crop-related attributes to generate accurate and actionable recommendations for farmers and other stakeholders.**

The solution is deployed using a lightweight and scalable Flask-based web application, ensuring broad accessibility across devices without requiring local software installation. Through the integration of predictive modelling, real-time analysis, and user-friendly visual outputs, the system seeks to enhance farm management, minimize uncertainty, and promote data-informed agricultural decision-making. Experimental results indicate that integrating multiple advisory modules improves the reliability of recommendations and offers a useful resource for farmers, researchers, and policymakers. The developed prototype also provides a strong basis for future advancements, including precision farming applications and real-time monitoring using IoT technologies.

***Key words*— Crop Recommendation, Fertilizer Recommendation, Crop yield prediction, Crop Price Prediction, logistic regression, soil, recommendation.**

I. INTRODUCTION

Agriculture plays a vital role in the global economy; however, it continues to encounter significant challenges arising from changing climatic patterns, limited natural resources, unstable commodity prices, and variations in soil and environmental characteristics [1][2]. Conventional agricultural decision-making is often an experience-based process that relies on past practices, which may be insufficient to address the increasing complexity of present-day farming systems. With the rapid growth of data availability, machine learning (ML) techniques have gained prominence as effective tools for converting large volumes of agricultural data into meaningful and actionable knowledge [3]. ML-enabled advisory solutions assist farmers in optimizing crop choices, managing inputs efficiently, predicting yields, and planning market participation, thereby enhancing productivity while promoting sustainability [4][5].

This study presents the design and implementation of an integrated and deployable platform entitled "Smart Agriculture through Machine Learning: A Multi-Aspect Farm Advisory System." The proposed framework combines four interrelated functional modules that support critical agricultural decisions:

1. Crop Suitability Analysis – Identifies appropriate crops for specific regions by analysing soil composition, weather conditions, and environmental factors.
2. Yield Estimation Module – Predicts the potential production of selected crops for a given location, enabling improved planning and risk assessment.
3. Market Price Forecasting – Anticipates future price movements to support informed decisions related to crop sales and storage strategies.
4. Nutrient Management Advisory – Suggests suitable fertilizer types and application levels based on soil nutrient status and crop-specific requirements.

All modules are deployed within a Flask-based web application that provides a simple and accessible interface through standard web browsers [6][7][8]. The platform accepts user inputs, executes trained machine learning models, and presents outputs through clear visualizations and recommendation reports. Owing to its modular architecture, the system supports scalability and can be adapted for real-time operation across diverse agricultural environments.

By effectively combining domain knowledge in agriculture with data-driven learning approaches, the proposed system seeks to advance sustainable and intelligent farming practices [6][9]. The platform reduces the complexity of predictive analysis while delivering timely and accurate guidance to farmers, thereby improving productivity and economic stability. Furthermore, the architecture establishes a foundation for future enhancements, including integration with IoT devices, remote sensing data, and precision agriculture technologies to develop a comprehensive smart farming ecosystem.

II. DATASET

The Smart Agriculture advisory framework is constructed using a combination of datasets obtained from open-access sources, official government platforms, and synthetically generated records. Together, these data sources capture essential information related to soil properties, weather conditions, crop attributes, production outcomes, and agricultural market behavior. Each dataset is utilized by a specific module within the system, as described below.

1. Crop Recommendation Dataset

The crop recommendation component employs a publicly available dataset obtained from the Kaggle data repository. This dataset is designed to support crop suitability analysis by providing detailed soil and environmental measurements relevant to agricultural land assessment [10]. The recorded attributes include soil nutrient concentrations and climatic indicators, namely nitrogen, phosphorus, and potassium levels measured in milligrams per kilogram, ambient temperature expressed in degrees Celsius, relative humidity percentage, soil pH value, and average rainfall measured in millimetres. The output variable corresponds to the recommended crop category, comprising twenty-two distinct crop classes.

Owing to its well-organized structure and comprehensive feature set, this dataset has been extensively adopted in machine learning–based agricultural research and serves as a reliable foundation for crop recommendation tasks.

The below table provides the features used for building the ML model.

	N	P	K	temperature	humidity	ph	rainfall	label
0	90	42	43	20.879744	82.002744	6.502985	202.935536	rice
1	85	58	41	21.770462	80.319644	7.038096	226.655537	rice
2	60	55	44	23.004459	82.320763	7.840207	263.964248	rice
3	74	35	40	26.491096	80.158363	6.980401	242.864034	rice
4	78	42	42	20.130175	81.604873	7.628473	262.717340	rice

Fig. 1 Crop Recommendation Dataset

2. Fertilizer Recommendation Dataset

The fertilizer advisory module is developed using an openly accessible dataset sourced from the Kaggle platform, commonly referred to as the Fertilizer Prediction Dataset. This dataset supports the identification of appropriate fertilizer options by analysing the relationship between soil nutrient status, environmental conditions, and crop-specific requirements [15].

The dataset consists of multiple input features, including ambient temperature and relative humidity, soil moisture content, and concentrations of essential macronutrients—nitrogen, phosphorus, and potassium. In addition, information regarding the cultivated crop type is included to ensure that fertilizer recommendations are aligned with crop nutrient demands. The corresponding output variable represents the recommended fertilizer category, encompassing both organic and inorganic fertilizer types.

By incorporating agronomic and environmental parameters, this dataset enables the development of an effective fertilizer recommendation model that assists farmers in selecting nutrient inputs tailored to specific soil and crop conditions.

	Temparature	Humidity	Moisture	Soil Type	Crop Type	Nitrogen	Potassium	Phosphorous	Fertilizer Name
0	26	52	38	Sandy	Maize	37	0	0	Urea
1	29	52	45	Loamy	Sugarcane	12	0	36	DAP
2	34	65	62	Black	Cotton	7	9	30	14-35-14
3	32	62	34	Red	Tobacco	22	0	20	28-28
4	28	54	46	Clayey	Paddy	35	0	0	Urea

Fig. 2 Fertilizer Recommendation Dataset

3. Crop Yield Dataset

The crop yield prediction module utilizes a combination of datasets collected from public data repositories and official government sources to ensure reliability and real-world relevance. Primary data are obtained from an open-access crop yield dataset available on the Kaggle platform, supplemented with information published by national agricultural and meteorological agencies. These sources include AgmarkNet, which provides insights into crop supply and market-related trends, the Indian Meteorological Department (IMD) [11] for weather-related parameters such as rainfall and temperature, and the Indian Council for Agricultural Research (ICAR) for region-wise crop productivity statistics.

The consolidated dataset captures a diverse set of attributes describing geographical, temporal, agronomic, and environmental factors. These include the state and district identifiers, cultivation year, growing season, and crop type, along with climatic indicators such as temperature, humidity, and soil moisture. Additional variables representing the cultivated area are incorporated to improve yield estimation accuracy. The target variable corresponds to total crop production, which serves as the basis for training and evaluating the yield prediction model.

By integrating multi-source data, the dataset enables robust modelling of yield variations across regions and seasons, supporting informed agricultural planning and risk assessment.

State_Name	District_Name	Crop_Year	Season	Crop	Temperature	humidity	soil moisture	area	Production
Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Arecanut	36	35	45	1254	2000
Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pulses	37	40	46	2	1
Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice	36	41	50	102	321
Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana	37	42	55	176	641
Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashewnut	36	40	54	720	165
Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Coconut	34	45	52	18168	13907
Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Dry ginger	34	55	62	36	100
Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sugarcane	35	50	59	1	2
Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sweet potato	25	55	55	5	15
Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Tapioca	36	35	45	40	169
Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Arecanut	37	40	46	1254	2061
Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Other Kharif pulses	36	41	50	2	1
Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Rice	37	42	55	83	300
Andaman and Nicobar Islands	NICOBARS	2001	Whole Year	Cashewnut	36	40	54	719	192

Fig. 3 Crop Yield Dataset

4. Market Price Dataset

The market price forecasting component is developed using price-related data collected from official agricultural information systems and open data platforms. The primary source of pricing and arrival statistics is AgmarkNet, which provides detailed records of commodity prices across various markets. Additional inputs are obtained from State Agricultural Marketing Boards and Agricultural Produce Market Committee (APMC) commodity bulletins to enhance regional coverage and data consistency.

For effective modeling, price information is organized on a crop-wise basis, with individual datasets maintained for each commodity under analysis. Each file contains time-stamped pricing attributes that capture market variability, including minimum traded price, maximum traded price, and modal price for a given date. The corresponding price date serves as the temporal reference for trend analysis and forecasting.

This structured, multi-source dataset enables accurate modeling of price fluctuations and supports informed decision-making related to crop marketing, storage, and sales strategies.

minprice	maxprice	modalprice	pricedate
3600	3900	3750	23-Apr-24
3700	4200	3950	23-Apr-24
3600	3900	3750	22-Apr-24
2091	2926	2480	22-Apr-24
3000	3500	3200	22-Apr-24
3200	3900	3550	22-Apr-24
1289	5269	2617	22-Apr-24
2169	5610	2667	22-Apr-24

Fig. 4 Crop Price Dataset

III. METHODOLOGY

The Smart Agriculture through Machine Learning system is implemented as a unified decision-support architecture composed of multiple interconnected modules that provide end-to-end farm advisory services. The methodological approach is structured as a sequential pipeline that begins with data collection and proceeds through data cleaning and transformation, feature extraction, model training, performance validation, and final deployment. This structured workflow is designed to support modular implementation, ease of scalability, and practical deployment in real-world agricultural environments.

1. System Architecture Overview

The proposed system architecture consists of four distinct machine learning components that operate independently while remaining functionally integrated within a single framework. These components include a crop recommendation unit, a crop yield estimation unit, a market price forecasting unit, and a fertilizer advisory unit. Each module is developed and trained using datasets and algorithms tailored to its specific agricultural objective.

A centralized web application built using the Flask framework serves as the common interface through which users can access all advisory functionalities. The modular structure of the system supports flexibility by allowing individual models to be updated, retrained, or replaced without disrupting the operation of the remaining components or the overall platform.

2. Data Collection and Integration

Data required for the proposed system are obtained from a combination of open Kaggle repositories, official agricultural data portals, and synthetically generated sources to achieve comprehensive representation of soil properties, climatic conditions, production statistics, and market behavior. As these datasets are derived from diverse origins, a data harmonization process is applied to unify file formats, standardize attribute names, and normalize measurement units. Market price information is organized as crop-wise time-series datasets, whereas soil characteristics and yield-related data are arranged in structured tabular formats to support efficient processing and modeling.

3. Data Preprocessing and Feature Engineering

Data preprocessing plays a vital role in improving the stability and predictive capability of the machine learning models used in the proposed system. A set of common preprocessing procedures is applied across all datasets to ensure data quality and consistency. These procedures include eliminating incomplete or contradictory entries, identifying and treating anomalous values through statistical boundary methods, and converting categorical attributes—such as crop category, growing season, and geographical region—into machine-readable formats. In addition, numerical features are scaled using normalization or standardization techniques when required to support effective model training.

In the context of crop yield prediction, the response variable representing production levels demonstrates noticeable skewness caused by differences in regional conditions and climatic factors. To reduce variance and facilitate improved model convergence, a logarithmic transformation of the form log(1 + production) is employed.

For crop price forecasting, historical price records are reformulated into a supervised learning structure by extracting time-dependent features. These include a continuous representation of time measured as the number of days from a reference point, as well as seasonal indicators generated using sine and cosine transformations of the day-of-year. The inclusion of these temporal features enables the models to effectively learn both long-term price movements and recurring seasonal patterns

IV. MACHINE LEARNING MODEL DEVELOPMENT

Each advisory component within the proposed framework applies a machine learning technique that aligns with the nature of the corresponding prediction task.

4.1 Crop Recommendation

The crop recommendation task is modeled as a multi-class classification problem, where soil nutrient concentrations (nitrogen, phosphorus, and potassium), along with environmental variables such as temperature, humidity, soil pH, and rainfall, are used as input features. Logistic Regression is selected for this module owing to its transparent decision boundaries, ease of interpretation, and consistent performance when applied to structured agricultural datasets.

4.2 Fertilizer Recommendation

The fertilizer advisory component aims to determine the most appropriate fertilizer option by analyzing soil moisture content, nutrient availability, prevailing climatic conditions, and crop type. A supervised classification framework is employed to learn the relationship between soil–crop characteristics and suitable fertilizer categories, enabling targeted nutrient recommendations.

4.3 Crop Yield Prediction

Yield estimation is addressed as a regression problem, utilizing predictors that include environmental conditions, cultivated land area, crop category, and seasonal attributes. A Random Forest Regressor is adopted for this module due to its robustness in modeling non-linear interactions and its effectiveness in handling high-dimensional, multivariate agricultural data.

4.4 Crop Price Prediction
Price forecasting is formulated as a time-dependent regression problem. Historical minimum and maximum market prices are aggregated to derive an average price, which serves as the response variable. The Random Forest Regressor is combined with engineered temporal and seasonal features to capture complex market behavior while avoiding the limitations of simple linear trend extrapolation. The trained model generates price predictions for a forward horizon of twelve months from the current reference date.

5. Model Training and Evaluation
For each advisory module, the corresponding dataset is partitioned into training and validation subsets using an 80:20 split. This division allows the models to learn underlying patterns from the majority of the data while reserving a portion for independent performance assessment. Evaluation criteria are selected based on the nature of the prediction task. Classification-based modules, including crop and fertilizer recommendation, are assessed using metrics such as accuracy, precision, and recall. In contrast, regression-oriented modules for yield estimation and market price forecasting are evaluated using mean squared error (MSE) and the coefficient of determination (R^2).
The adoption of task-specific evaluation metrics enables a comprehensive assessment of both predictive reliability and the ability of the models to generalize effectively to previously unseen data.

6. Deployment and User Interface
The trained machine learning models are integrated into a web application developed using the Flask framework, offering a lightweight and user-friendly interface for end users. Through this interface, users can provide inputs related to soil characteristics, prevailing weather conditions, crop information, and regional attributes. The system processes these inputs in real time and displays the resulting predictions in an easily interpretable format, including descriptive textual recommendations and interactive visual elements such as time-series charts for market price analysis.
This web-based deployment approach ensures cross-platform accessibility and supports real-time model inference without the need for dedicated hardware resources or specialized software installations, making the system practical for widespread adoption.

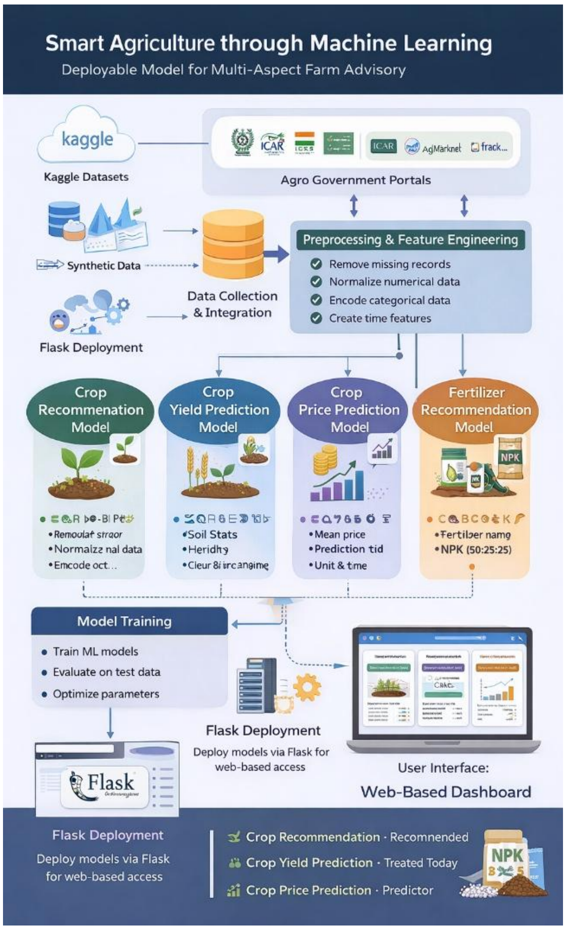

Fig 5: Overall Architecture

V. RESULTS AND CONCLUSION

The performance of the proposed Smart Agriculture through Machine Learning framework was assessed using four integrated advisory components, namely crop suitability analysis, fertilizer recommendation, market price forecasting, and crop yield estimation. The complete system was deployed as an interactive web application developed with the Flask framework, allowing users to obtain predictions and visualize outputs in real time. Experimental outcomes indicate that the proposed approach effectively delivers reliable, transparent, and practical agricultural recommendations, demonstrating its potential to support informed decision-making in modern farming practices.

5.1 Crop Recommendation Results
The crop recommendation module leverages inputs such as soil nutrient concentrations (nitrogen,

phosphorus, potassium), temperature, humidity, rainfall, and pH to determine the most appropriate crops for specific field conditions. The system's web interface displays crop suggestions along with suitability indicators corresponding to the parameters provided by the user.

Evaluation shows that the model consistently generates agronomically meaningful recommendations, demonstrating its ability to effectively map soil and climatic factors to crop selection. Recommendations are presented through a user-friendly interface designed for farmers, facilitating ease of interpretation and supporting practical decision-making in the field.

5.2 Fertilizer Recommendation Results

The fertilizer advisory module utilizes features such as soil moisture, nutrient content, temperature, humidity, soil classification, and crop type to identify the most suitable fertilizer for a given field. Evaluation results show that the model reliably maps soil and crop characteristics to appropriate fertilizer categories, including options such as DAP, Urea, and NPK blends.

Beyond identifying the fertilizer type, the system also provides practical application guidance, including optimal timing and suitability for specific crops, enhancing the clarity and usability of the recommendations. This module illustrates how machine learning can contribute to precision agriculture by optimizing fertilizer application, reducing overuse, and supporting sustainable crop management.

5.3 Crop Price Prediction Results

The crop price forecasting module is formulated as a time-series regression task, employing a Random Forest Regressor augmented with engineered temporal and seasonal features. Historical market price data serve as the basis for predicting price trends over a 12-month horizon from the current date. Evaluation indicates that the model effectively captures both non-linear market fluctuations and seasonal variations, avoiding simplistic linear extrapolation. Predicted prices are presented through an interactive time-series visualization that highlights the current price, projected price trajectory, and the maximum predicted price marked with a distinct star indicator. The model demonstrated robust generalization, achieving an R^2 score of approximately 0.71 on the test dataset, confirming its reliability for practical agricultural market forecasting.

5.4 Crop Yield Prediction Results

The crop yield estimation module predicts expected production by considering factors such as region, season, crop type, temperature, humidity, soil moisture, and cultivated area. To address skewness and variability in yield data, a Random Forest Regressor was trained on logarithmically transformed production values.

The system provides quantitative yield forecasts expressed in quintals and complements these predictions with practical agricultural recommendations, including guidance on irrigation management, temperature regulation, and seasonal planning. Evaluation demonstrates that the model successfully captures intricate relationships between environmental and agronomic variables, delivering reliable predictions with strong generalization to unseen data.

5.5 System Deployment and Usability

All four advisory modules are combined within a single web application developed using the Flask framework, offering a cohesive dashboard for smart agriculture recommendations. The interface enables users to easily switch between modules, provide input parameters, and access predictive outputs in real time. The successful implementation highlights the system's practical utility and demonstrates its scalability for deployment in real-world agricultural settings.

VI. CONCLUSION

This study introduced an integrated smart agriculture advisory platform that employs machine learning to tackle multiple decision-making challenges within a single deployable framework. Unlike conventional approaches that address individual prediction tasks in isolation, the proposed system consolidates crop recommendation, fertilizer guidance, yield estimation, and market price forecasting into a unified, web-accessible solution.

Experimental evaluation demonstrates that the chosen models, particularly the Random Forest Regressor for yield and price forecasting, provide robust predictive performance while capturing non-linear relationships and accommodating variability in real-world agricultural data. Techniques such as feature engineering, logarithmic transformation of target variables, and seasonal encoding contributed to improved accuracy and model stability.

The Flask-based deployment ensures an intuitive interface that is both accessible and interpretable for farmers and agricultural stakeholders. Overall, the system shows strong potential for supporting data-driven farm management, optimizing resource use, and enhancing productivity. Future enhancements could include integration with real-time weather data APIs, satellite imagery analysis, deep learning-based forecasting, and mobile application deployment to broaden system accessibility and functionality.

REFERENCES

[1] Tawseef Ayoub Shaikh, Tabasum Rasool, Faisal Rasheed Lone. "Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming." *Computers and Electronics in Agriculture,* Volume 198, July 2022, 107119 Internet: https://doi.org/10.1016/j.compag.2022.107119 [referred in section 1.1]

[2] Dhruvi Gosai, Chintal Raval, Rikin Nayak, Hardik Jayswal, Axat Patel. "Crop Recommendation System using Machine Learning." *International Journal of Scientific Research in Computer Science, Engineering and Information Technology,* Volume 7, Issue 3 Page Number: 554-557, May-June-2021 [referred in section 2.1]

[3] Rachid Ed-daoudi1, Altaf Alaoui2 , Badia Ettaki3 , Jamal Zerouaoui. "A Predictive Approach to Improving Agricultural Productivity in Morocco through Crop Recommendations." *International Journal of Advanced Computer Science and Applications,* Vol. 14, No. 3, 2023[referred in section 2.2]

[4] Ankit R. Sawant, Yash Sivramkrishnan, Hemish Ganatra, Ameya A. Kadam. "Smart Crop-Precision Agriculture using Machine Learning." *International Journal of Research in Engineering, Science and Management,* Volume 6, Issue 10, October 2023[referred in section 2.3]

[5] Devdatta A, Mr.Santosh Mahagaonkar. "Prediction of crop yield and fertilizer Recommendations using ML Algorithms." *International Journal of Engineering Applied Sciences and Technology*, 2019 Vol. 4, Issue 5[referred in section 2.4]

[6] Navod Neranjan Thilakarathne, Muhammad Saifullah Abu Bakar, Pg Emerolylariffion Abas, Hayati Yassin. "A Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming." Internet: https://doi.org/10.3390/su12219138, 2020 [referred in section 2.5]

[7] R. JOHN MARTIN et.al., "XAI-Powered_Smart_Agriculture_Framework_for_Enhancing_Food_Productivity_and_Sustainability",IEEE Research Article, Sep 2024

[8] Vedika Sanjay et.al.," An AI-Driven Smart Crop Recommendation and Advisory Framework", International Research Journal on Advance Engineering Hub, July 2025

[9] Mahmoud Y. Shams et.al., "Enhancing crop recommendation systems with explainable artificial intelligence a study on agricultural decision-making", Neural Computing and Applications, Jan 2024

[10] Sourabh Shastri et.al., "Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence", Scientific Reports, July 2025

[11] PRITESH PATIL et.al., "Crop Selection and Yield Prediction using Machine", Agriculture Journal, May 2023

[12] MD Shaifullah Sharafat et.al., "An IoT-enabled AI system for real-time crop prediction using soil and weather data in precision agriculture", Science Direct, July 2025

[13] Tanmay Thorat et.al., "Intelligent insecticide and fertilizer recommendation system based on TPF-CNN for smart farming" Science Direct', Feb 2023

[14] Rahul S. Pachade and Dr. Avinash Sharma, "Machine learning for weather-specific crop recommendation", International Journal for Health Sciences, May 2022

[15] Jyoti Deone,Dr. Khan Rahat Afreen,Rupali Mangrule , "Machine Learning Approaches for Crop Prediction and Fertilizer Recommendation based on Soil Nutrients", Journal of Electrical Systems, May 2024

[16] Kranti G. Sapkal and Avinash B. Kadam," Predictive Modeling Approach for Crop Recommendation Using Machine Learning Algorithms" ,EPJ Web of Conferences, June 2025