AgroAdvisor: Intelligent Agriculture Decision Support System using ML

[1]Chetana Patil, [2]Revati Dakhare, [3]Mohit Pawar, [4]Aavez Sayyad

[1]Assistant Professor,[2]Student, [3]Student, [4]Student

Department of Computer Engineering,

Dhole Patil College of Engineering, Wagholi, Pune- 412207, India

Abstract: Farming is one of the major segments that influences a country's financial development. However, farmers face numerous challenges such as unpredictable weather patterns, soil degradation, and resource constraints. Modernising agriculture through technology is crucial for meeting global food demands sustainably. This project "AgroAdvisor: Intelligent Agriculture Decision Support System using ML" introduces an innovative approach that leverages data analytics and machine learning to provide precise fertilizer recommendations and crop prediction based on soil nutrient content. By optimizing the balance of nitrogen, potassium, and phosphorus, the project aims to enhance crop yield while minimizing environmental impact. By integrating ML algorithms such as XGBoost and Random Forest, Agroadvisor analyzes soil content, weather forecasts, and historical crop performance to provide accurate predictions and recommendations for crop selection, fertilizer application, and weather-related decisions. The proposed web application offers farmers user-friendly access to tailored guidance, revolutionizing nutrient management practices and fostering more efficient and eco-friendly agriculture. The Intelligent Agriculture Decision Support System showcases the transformative capacity of machine learning in revolutionizing agricultural methodologies. By combining soil and weather insights with advanced predictive modelling, the project empowers farmers to optimize crop yields, enhance resource efficiency, and embrace sustainable agricultural methods.

Index Terms – Machine Learning, Crop Prediction, Fertilizer Recommendation, XGBoost, Randomforest, Weather Forcast.

I. INTRODUCTION

The agricultural sector is a cornerstone of global food production, providing the sustenance required for a burgeoning global population. However, it encounters numerous challenges that require innovative solutions. One of the primary dilemmas is the perennial conundrum of crop selection, wherein farmers must decide which crops to cultivate based on factors like soil quality, climate conditions, and market demand. This decision significantly impacts crop yield, resource utilization, and the environment. Furthermore, optimizing the use of fertilizers is a critical aspect of modern farming. Excessive or inadequate fertilizer application can lead to diminished yields, increased costs, and adverse environmental consequences.

Historically, farmers have relied on traditional knowledge and localized practices to navigate these complexities. However, these approaches often fall short in an era marked by dynamic environmental shifts and increasingly unpredictable weather patterns. The emergence of data science, machine learning, and artificial intelligence has provided a chance to transform decision-making processes in agriculture.

By harnessing the power of these technologies, it is possible to develop intelligent agricultural decision support systems that provide farmers with real-time, data-driven insights.

In this context, "AgroAdvisor" represents a significant leap forward in the evolution of agriculture. By integrating machine learning and deep learning algorithms, this project seeks to empower farmers with a comprehensive decision support system. It endeavours to assist them in determining the most suitable crops to plant, based on a nuanced analysis of their unique land conditions. Moreover, it aims to optimize fertilizer recommendations, ensuring that nutrients are managed effectively, and environmental impact is minimized. Beyond these technical aspects, the project emphasizes the need for user-friendliness, providing easily interpretable recommendations and insights that can be seamlessly integrated into the farming process.

The primary objectives of this project are outlined as follows:

1) Crop recommendation
2) Recommendation of Fertilizer for soil improvement
3) Weather forecast for up to 5 days

II. PROBLEM STATEMENT

Farmers frequently encounter difficulties in deciding the most appropriate crops for cultivation and the best fertilization approaches for their fields, taking into consideration factors such as soil type, weather conditions, and market demand. Farmers need a comprehensive decision support system that leverages machine learning and deep learning to provide real-time insights and recommendations for making informed crop choices that are crucial for optimizing yield and profitability.

III. OBJECTIVES

- Crop Prediction: Develop accurate predictive models to forecast suitable crops based on soil composition, climate conditions, and historical data.
- Fertilizer Recommendation: Utilize machine learning algorithms to analyze soil nutrient levels and provide personalized recommendations for fertilizer application, optimizing crop nutrition and yield.
- Weather Forecasting: Integrate weather data analysis to deliver timely and accurate forecasts for at least 5 days, enabling farmers to plan agricultural activities effectively and mitigate weather-related risks.

IV. LITERATURE REVIEW

In recent years, the agricultural sector has witnessed a surge in research exploring the integration of advanced Machine Learning techniques. Several studies have delved into predicting crop outcomes, recommending suitable crops, and optimizing yield through various algorithms and data parameters.

For instance, Kevin Tom Thomas et al. [1] utilized KNN with cross-validation exclusively on soil parameters for crop prediction. Conversely, Mahendra N. et al. [2] employed Decision Tree algorithms considering both soil and weather parameters, with rainfall estimation derived from previous years' data using SVM. Similarly, Mansi Shinde et al. [3] and Kiran Shinde et al. [4] utilized Random Forest for crop recommendation and fertilizer suggestions, respectively, focusing on soil attributes.

Swapneel Chakraborty et al. [5] explored crop recommendation through Neural Networks, integrating soil parameters and environmental factors. Sonal Jain et al. [6], on the other hand, incorporated soil and weather parameters for crop selection, although omitting crucial soil metrics like NPK values. A. Suresh et al. [7] predicted crop yield for major crops in Tamilnadu using modified KNN and K-Mean, albeit neglecting other crops.

Various methodologies have been proposed to enhance crop selection and recommendation processes. A research paper [9] underscores the importance of considering government policies, market dynamics, and production rates in crop selection, advocating for a comprehensive Crop Selection Method (CSM) to boost net yield. Shilpa Pande et al. [10] developed a Crop Recommendation system based on Random Forest, incorporating location and soil properties to predict crop yield. Angu Raj et al. [11] explored IoT-driven Crop Recommendation, utilizing sensors for real-time data collection on soil parameters and applying Forest and Naïve Bayes algorithms for analysis.

T. Ragunthar et al. [12] tackled the Crop Selection problem by integrating the Apriori algorithm and Decision tree induction, considering attributes such as duration, water requirements, budget, soil type, and market factors. Sridhar Mhasikar et al. [13] explored Predicting Agricultural Cultivation using IoT, focusing on soil factors and employing Random Forest and Naïve Bayes algorithms.

Furthermore, Lakshmi N. et al. [14] proposed a comprehensive crop recommendation system leveraging big data to consider various soil attributes such as drainage, texture, and pH levels. M.V.R Vivek et al. [15] investigated the utilization of machine learning algorithms including Naive Bayes, multi-layer perceptron, JRIP, Jf48, and SVM based on meteorological data for crop recommendation. Hao Zhang et al. [16] explored the design of crop recommendation and fertilizer recommendation systems, integrating meteorological, crop, and soil data for analysis.

Table 1 Limitations of Revived Work

Authors	Technologies Used	Limitations
Kevin Tom Thomas et.al [1]	• Cross-validation is applied alongside KNN, and both Decision Tree and Naïve Bayes algorithms are utilized. • Parameters: Nitrogen, Phosphorous, Potassium and soil pH value.	Environmental factors are not considered in the process of recommending appropriate crops.
Mansi Shinde et.al [3]	• Random Forest, Naive Bayes, and the ID3 machine learning algorithms are employed. • Parameters: Nitrogen, Phosphorous and Potassium.	Only soil parameters are taken into consideration. Weather parameters are not taken into consideration.

Sonal Jain et.al [6]	• Random Forest algorithm is used. • Attributes include temperature, rainfall, soil type, soil pH, water capacity, and soil fertility.	Soil components such as Nitrogen, phosphorus, and Potassium are not factored into account.
A. Suresh et. al [7]	• KNN algorithm is used. • Parameters: Rainfall, groundwater, soil type and cultivated area.	Crop recommendation is done only for major crops i.e. Rice, Sugarcane, Maize, Raggi and Tapioca. Other crops are not taken into consideration.

V. PROPOSED OUTCOMES

Agriculture stands as the cornerstone of the nation's economy, serving as a vital income source. Soil type, pH levels, rainfall, and environmental conditions constitute pivotal elements upon which agriculture relies. Erroneous crop selection could result in diminished crop yield, imposing significant losses on farmers. Accessing appropriate crop recommendations prior to planting holds immense potential in augmenting crop output and facilitating informed decisions regarding storage and commercial endeavours. Soil inadequacies or lacking essential crop nutrients may hinder farmers from cultivating desired crops. By receiving recommendations for enhancing soil fertility, farmers can pursue desired crops effectively. Additionally, adverse weather conditions pose significant threats to crop production; hence, accessing weather forecasts for the upcoming five days aids farmers in undertaking necessary precautions.

1. Crop Prediction

Following the preprocessing and partitioning of the dataset into training and testing subsets, a range of Machine Learning algorithms were applied, including Decision Tree [17], Naïve Bayes [18], SVM [19], Linear Regression [20], Random Forest [21], and XGBoost [22]. Among these algorithms, XGBoost demonstrates the highest accuracy level.

XGBoost :

XGBoost, a gradient-boosting algorithm, accelerates the performance and efficiency of tree-based machine learning models, particularly decision trees. Initially, in the first iteration, XGBoost may exhibit lower accuracy due to the rudimentary nature of the model. However, as the number of iterations increases, the model refines its loss function optimization utilizing the Gradient Descent technique. This iterative process continues until the model reaches a threshold where further optimization of loss is not feasible. Consequently, the model's accuracy progressively improves with each iteration. In the proposed system, a training size of 0.8, corresponding to 80% of the dataset, is designated for training purposes. Subsequently, the trained model is evaluated on the remaining 20% of test data. At this juncture, the XGBoost model is trained using input parameters such as Nitrogen, Phosphorus, Potassium, pH, rainfall, temperature, and humidity, provided by the farmer. By employing the predict_proba() function, probabilities for each output class are determined. The top three values with the highest probabilities are then selected, indicating the recommended crops based on the given input features. As a result, three crops are suggested as the most suitable options for cultivation considering the provided input parameters.

2. Crop Recommendation Based on Production

Some farmers may lack data regarding the soil ingredients on their land. In such cases, the system will accept inputs such as State, District, and Season. Upon analyzing these inputs alongside the dataset, the system will suggest the top three crops with the highest production within the specified State, District, and Season.

3. Soil Improvements Recommendation

Soil fertility hinges on its NPK value. In the proposed system, inputs such as soil properties (Nitrogen, Potassium, Phosphorus) and the desired crop are required. Subsequently, the system analyzes these inputs and the dataset to provide recommendations on enhancing soil fertility for cultivating the desired crop.

4. Weather Forecast

Climate forecasting and prediction involve collecting and disseminating information regarding forthcoming weather conditions derived from weather forecasts. The farmer inputs the state and city, following which the system retrieves climate data using the Weather API [23]. This enables farmers to take necessary precautions and mitigate potential losses.

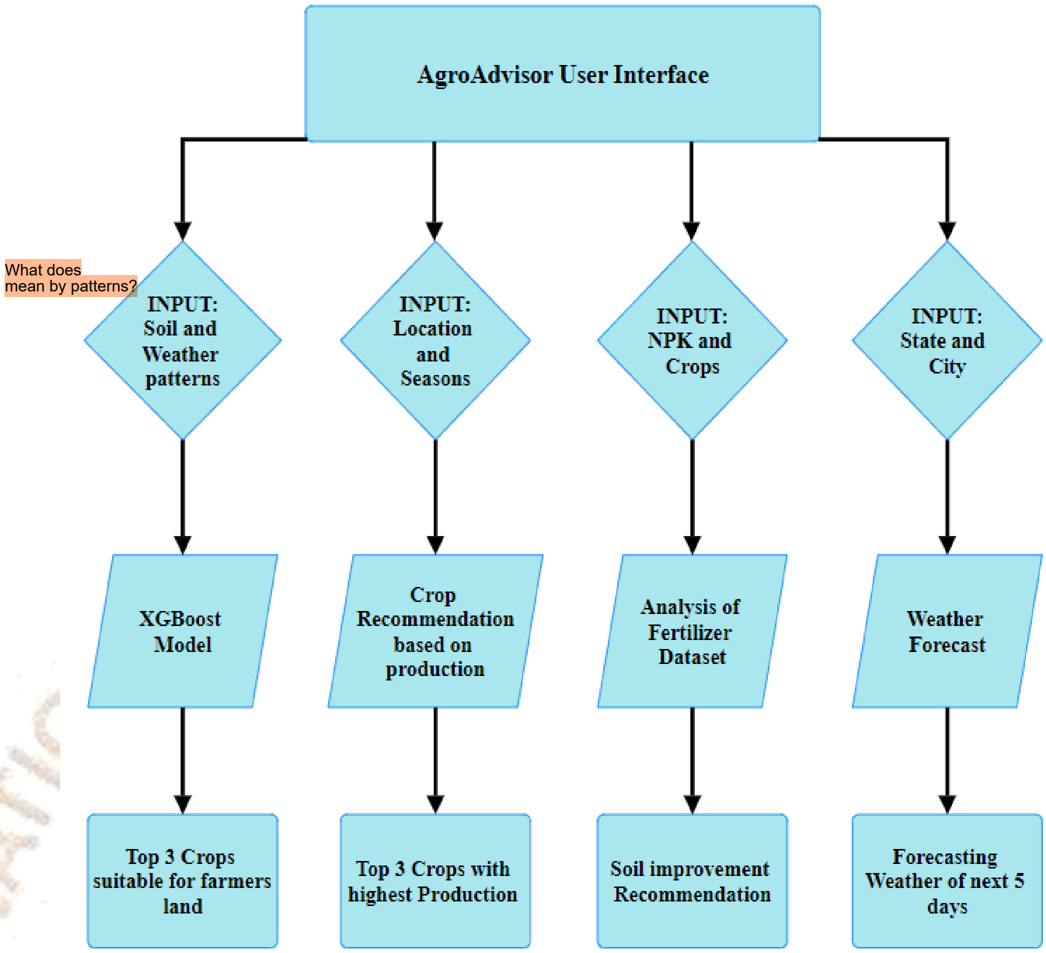


Fig.1 System Design

VI. RESULT AND DISCUSSION

This section encompasses the implementation and outcomes of all features incorporated into the proposed system. Training and testing any Machine Learning model necessitate a dataset. For the proposed crop recommendation system, the dataset was sourced from Kaggle [24], a renowned website. This dataset originates from soil testing labs and encompasses soil-specific attributes. Various crops like maize, rice, and coffee are considered within this system. The attributes encompass different nutrients such as nitrogen, potassium, phosphorus, and soil pH values. Additionally, it includes weather parameters like temperature, humidity, and rainfall.

Furthermore, another dataset utilized in this system is also sourced from Kaggle [25], encompassing attributes like Season, State name, Area, District name, and Production. Datasets acquired from open sources necessitate cleaning. Noisy, incomplete, and incompatible data are unsuitable for processing by Machine Learning algorithms and thus require cleaning.

1. Crop Prediction

Following dataset preprocessing, it was divided into an 80% training dataset and a 20% testing dataset. Various Machine Learning Algorithms including Decision Tree, Naïve Bayes, SVM, Logistic Regression, Random Forest, and XGBoost were applied to the dataset. Upon evaluating the model accuracies, the Decision Tree achieved 90%, Naïve Bayes attained 99%, SVM reached 10%, Logistic Regression achieved 95.22%, Random Forest scored 99%, and XGBoost obtained 99.3%.

The proposed system opts for the XGBoost algorithm to predict the top three optimal crops for farmers based on soil and weather parameters due to its highest accuracy.

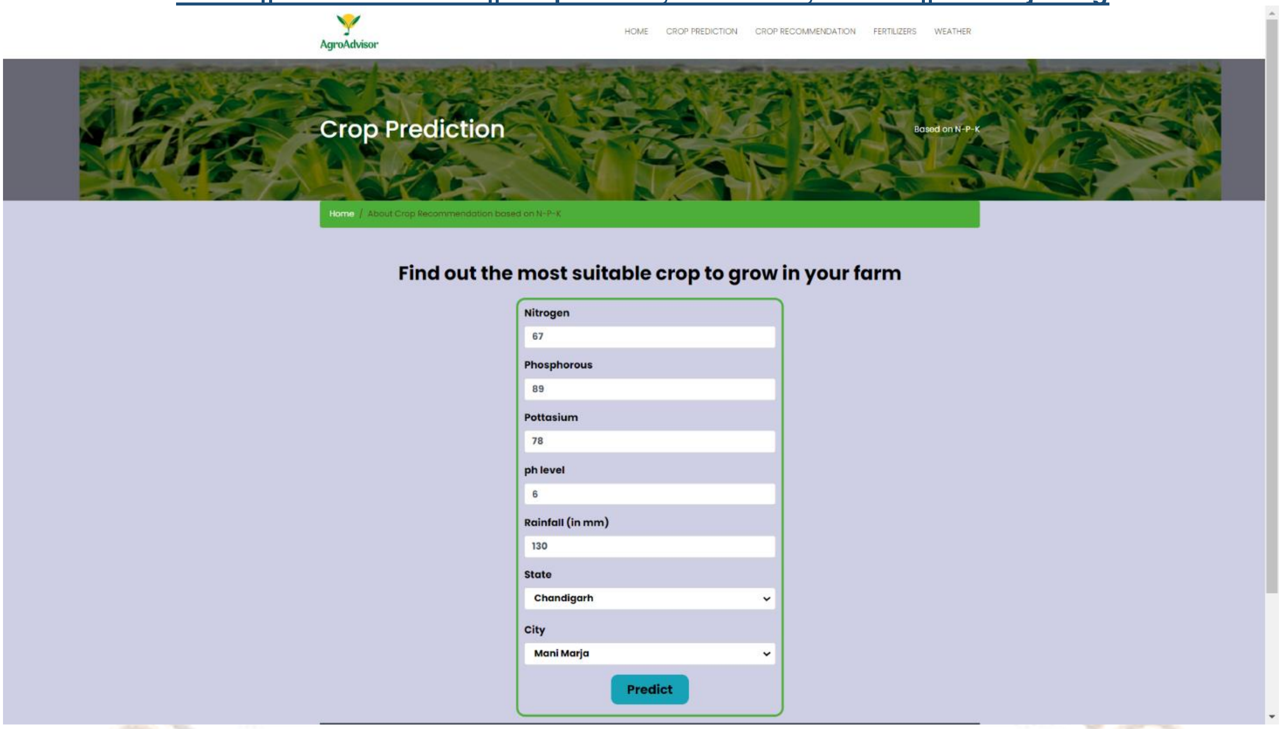


Fig.2 Crop Prediction Input

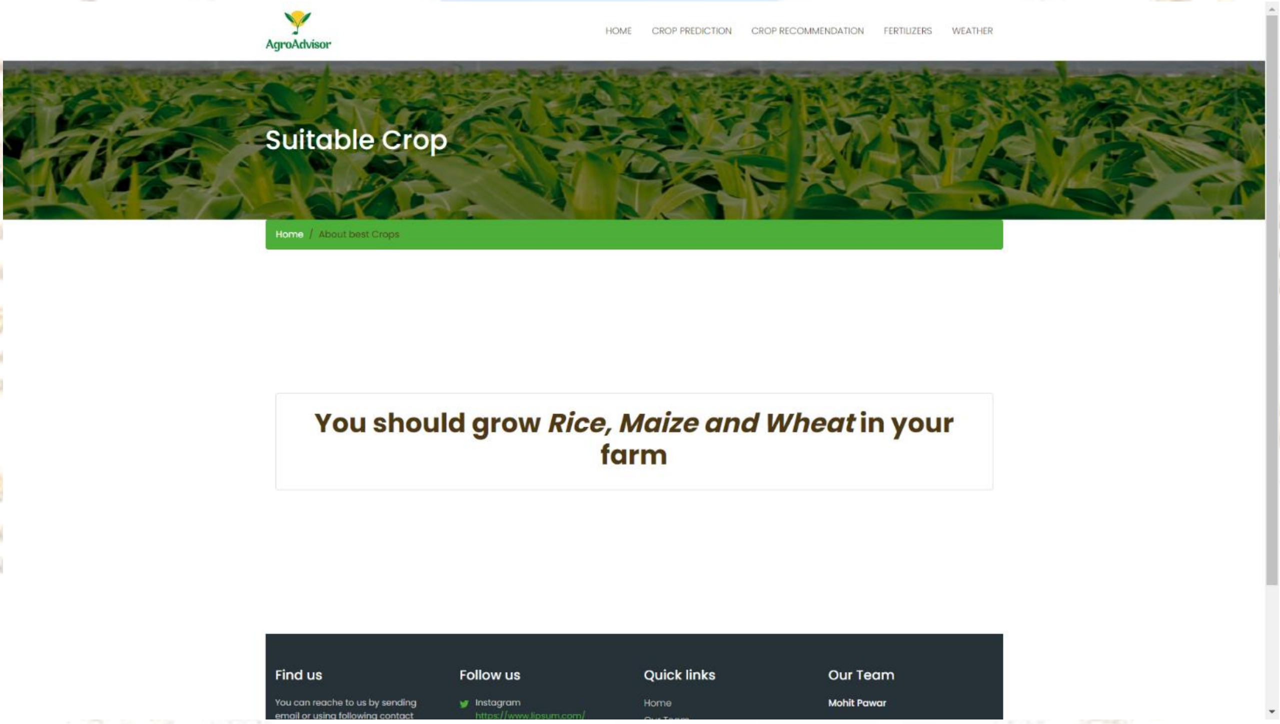


Fig.3 Crop Prediction Output

2. Crop Recommendation Based on Production

The Proposed system accepts inputs such as State, District, and Season, as depicted in FIG 4.2.

Following input acquisition, the system proceeds to suggest the top three crops with the highest production in the specified season and region, as illustrated.

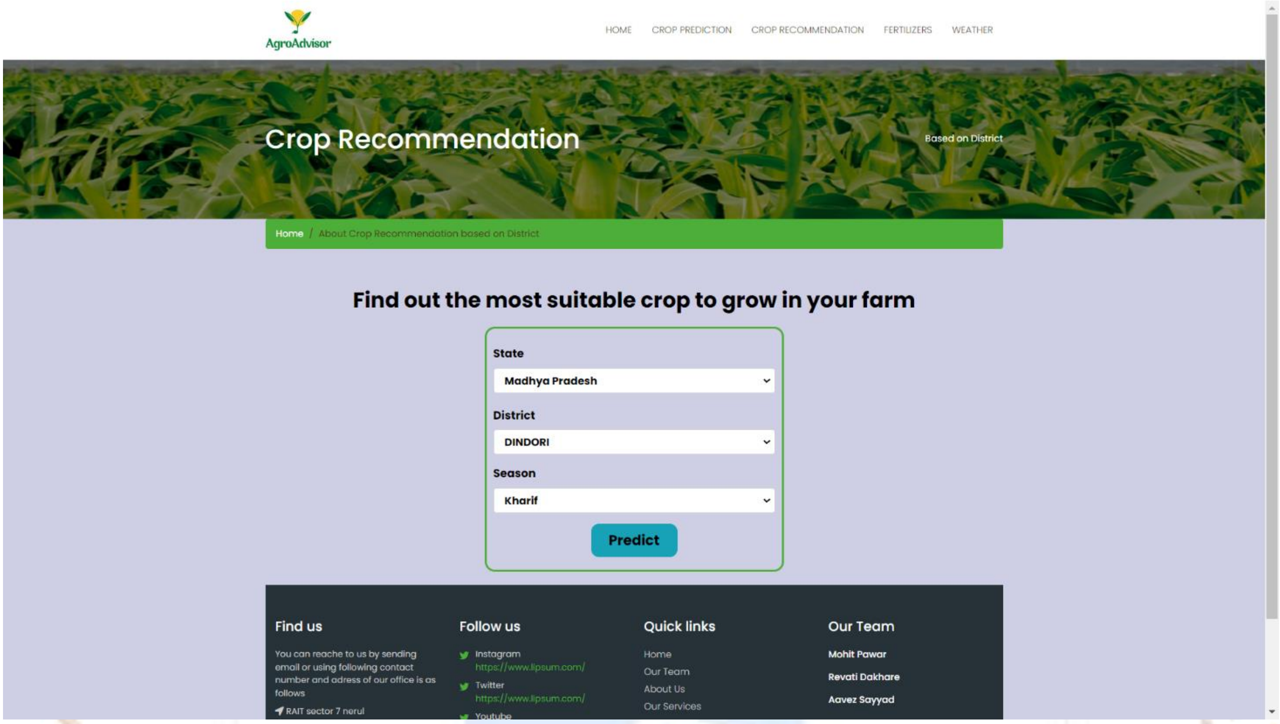


Fig.4 Crop Recommendation Input

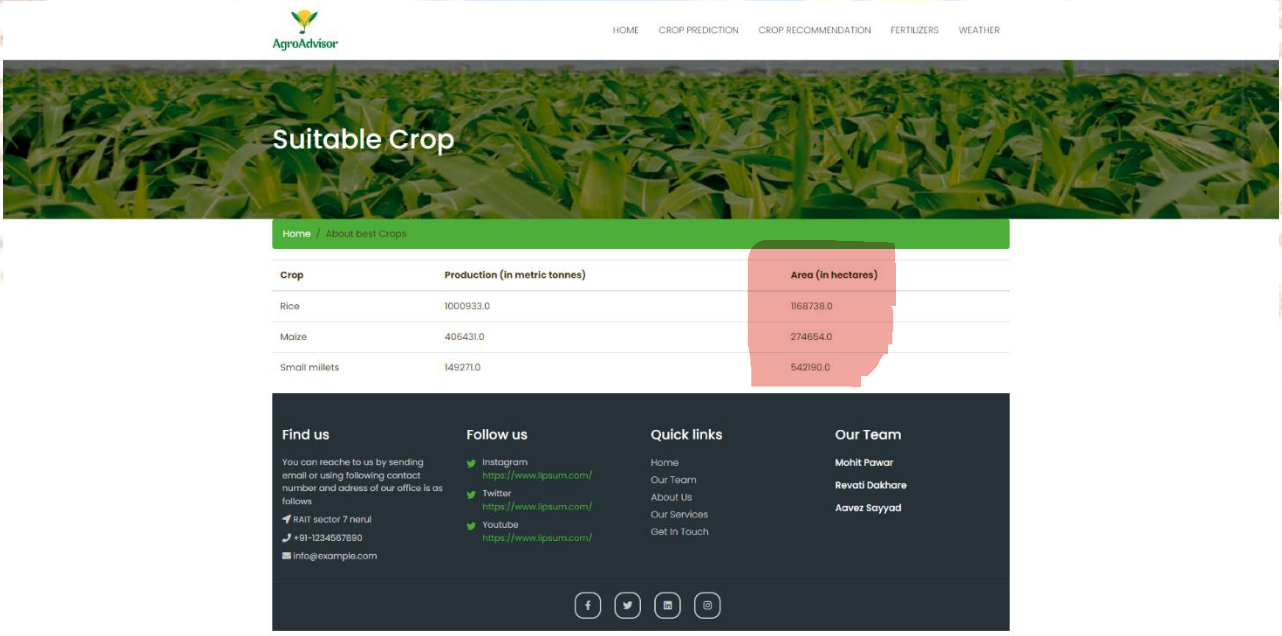


Fig.5 Crop Recommendation Output

3. Soil Fertility Improvement

The farmer provides input regarding the desired crop for cultivation and the NPK levels of the soil. Upon analyzing the dataset, the system identifies the soil's deficiencies or excesses in specific nutrients and recommends soil enhancements to render it conducive for cultivating the desired crop.

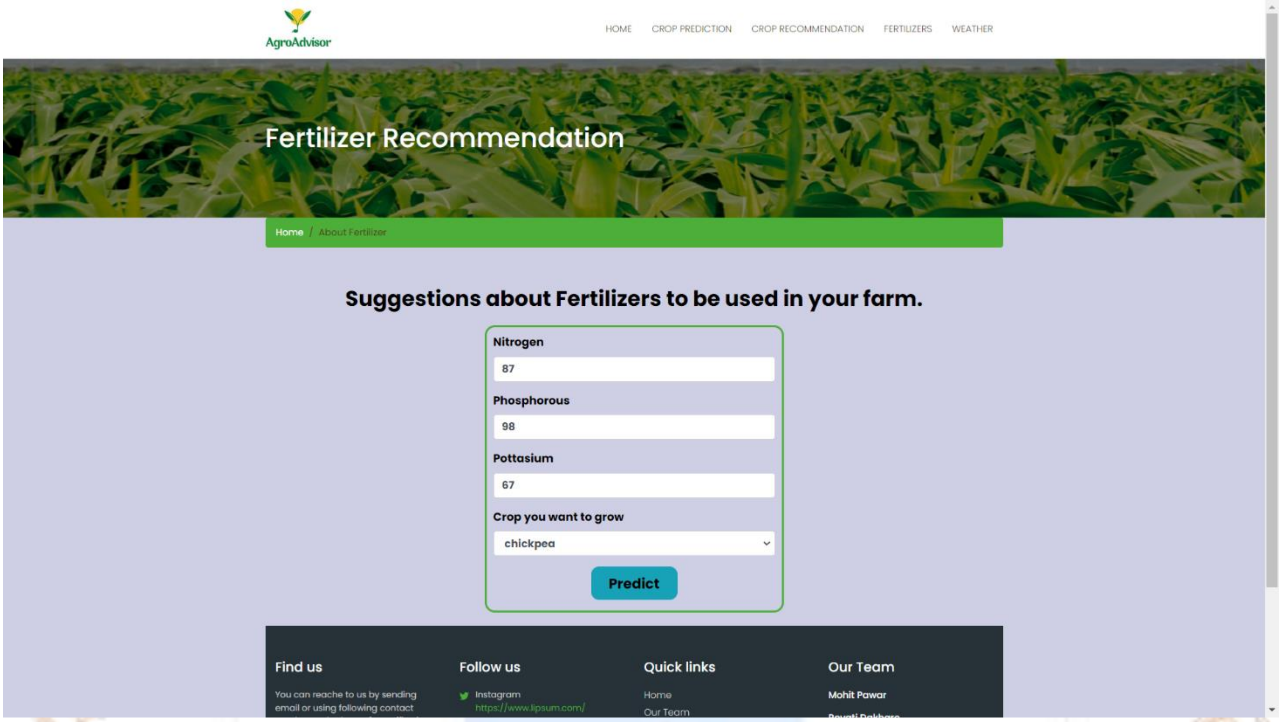


Fig.6 Fertilizer Recommendation Input

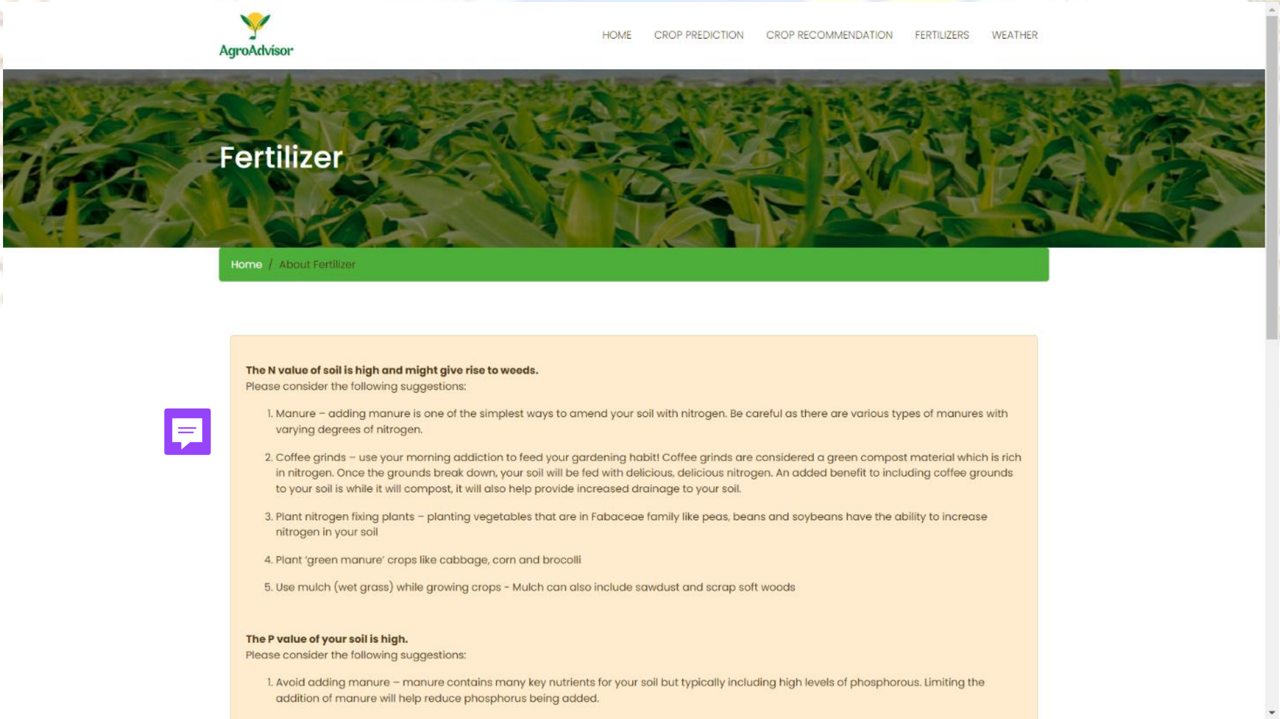


Fig.7 Fertilizer Recommendation Output

4. Weather Forecast

The farmer inputs the State and City, following which the system forecasts the weather for the next five days. The forecast includes parameters such as temperature, minimum temperature, maximum temperature, pressure, humidity, weather description, wind speed, wind gust, visibility, date, and time.

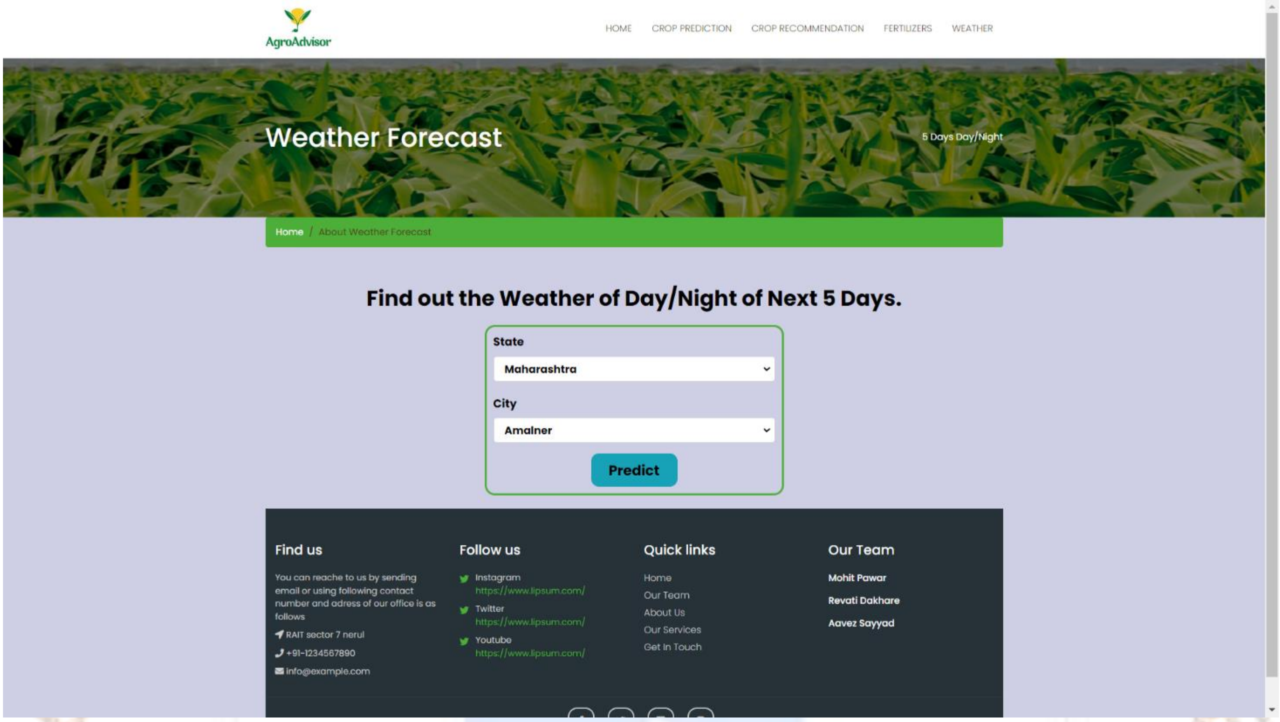
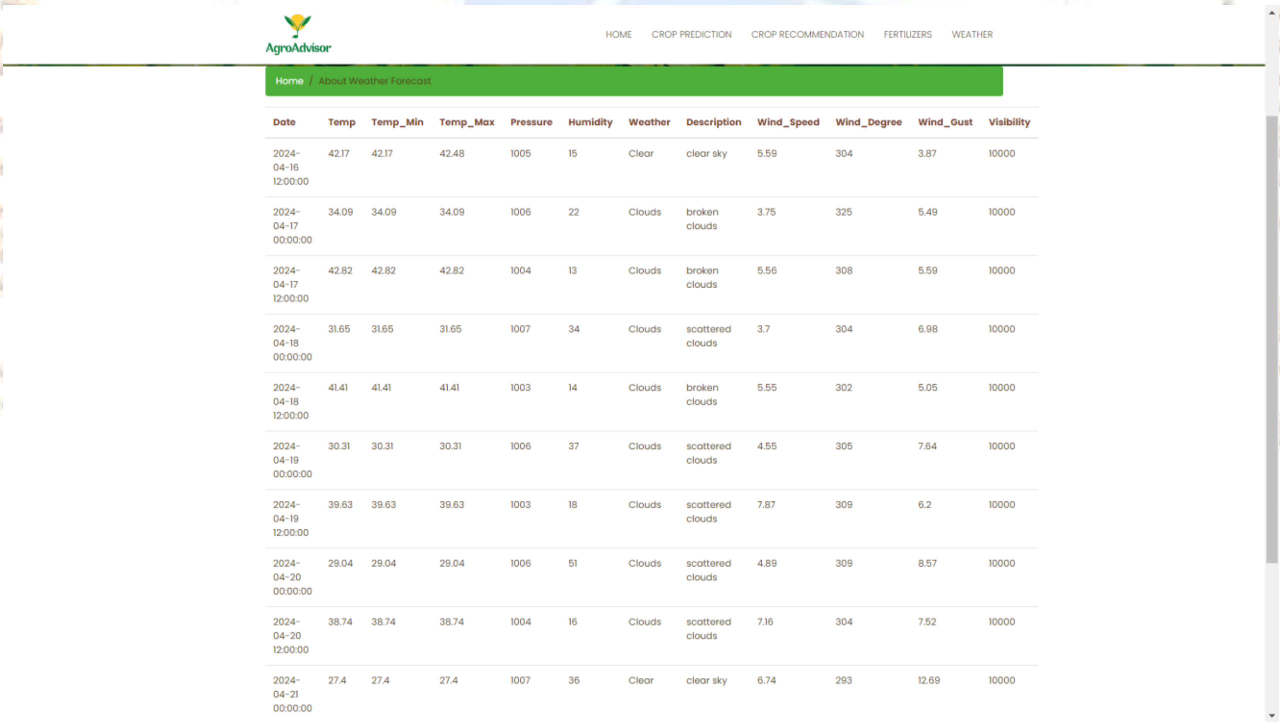


Fig.8 Weather Forecast Input



Date	Temp	Temp_Min	Temp_Max	Pressure	Humidity	Weather	Description	Wind_Speed	Wind_Degree	Wind_Gust	Visibility
2024-04-16 12:00:00	42.17	42.17	42.48	1005	15	Clear	clear sky	5.59	304	3.87	10000
2024-04-17 00:00:00	34.09	34.09	34.09	1006	22	Clouds	broken clouds	3.75	325	5.49	10000
2024-04-17 12:00:00	42.82	42.82	42.82	1004	13	Clouds	broken clouds	5.56	308	5.59	10000
2024-04-18 00:00:00	31.65	31.65	31.65	1007	34	Clouds	scattered clouds	3.7	304	6.98	10000
2024-04-18 12:00:00	41.41	41.41	41.41	1003	14	Clouds	broken clouds	5.55	302	5.05	10000
2024-04-19 00:00:00	30.31	30.31	30.31	1006	37	Clouds	scattered clouds	4.55	305	7.64	10000
2024-04-19 12:00:00	39.63	39.63	39.63	1003	18	Clouds	scattered clouds	7.87	309	6.2	10000
2024-04-20 00:00:00	29.04	29.04	29.04	1006	51	Clouds	scattered clouds	4.89	309	8.57	10000
2024-04-20 12:00:00	38.74	38.74	38.74	1004	16	Clouds	scattered clouds	7.16	304	7.52	10000
2024-04-21 00:00:00	27.4	27.4	27.4	1007	36	Clear	clear sky	6.74	293	12.69	10000

Fig.9 Weather Forecast Output

VII. CONCLUSION

In summary, the proposed system enables farmers to make informed decisions that result in increased profitability. By leveraging advanced machine learning techniques such as Random Forest, SVM, Decision Tree, Logistic Regression, Naive Bayes, and XGBoost, the system identifies the optimal classifier, with XGBoost demonstrating an impressive accuracy of 99.3%. This proficiency enables farmers to select the most lucrative crops for cultivation, aiding in strategic storage and business planning. Moreover, the system goes beyond crop selection by providing soil improvement recommendations, ensuring an ideal environment for desired crops. The integration of a weather forecast feature for the next five days equips farmers with valuable information to take necessary precautions. The promising results highlight the transformative potential of machine learning in revolutionizing agriculture. Looking ahead, the system can further evolve by incorporating Internet of Things (IoT) technology for real-time soil monitoring, exemplifying a smart and efficient approach to farming practices.

REFERENCES

[1] Kevin Tom Thomas, Varsha S, Merin Mary Saji, Lisha Varghese and Er. Jinu Thoma, “Crop Prediction Using Machine Learning”, International Journal of Future Generation Communication and Networking, Vol. 13, Issue: 3, pp. 1896-1901, 2020.

[2] Mahendra N, Dhanush Vishwakarma, Nischitha K, Ashwini, and Manjuraju M. R, “Crop Prediction using Machine Learning Approaches”, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH &TECHNOLOGY (IJERT) Volume 09, Issue 08 (August 2020).

[3] Mansi Shinde, Kimaya Ekbote, Sonali Ghorpade, Sanket Pawar and Shubhada Mone, "Crop Recommendation and Fertilizer Purchase System", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (2), 665- 667, 2016.

[4] Kiran Shinde, Jerrin Andrei and Amey Oke, "Web-Based Recommendation System for Farmers", International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169, Volume: 3 Issue: 3 1444 – 1448.

[5] P. A, S. Chakraborty, A. Kumar and O. R. Pooniwala, "Intelligent Crop Recommendation System using Machine Learning," 5th International Conference on Computing Methodologies and Communication (ICCMC), pp. 843-848, doi: 10.1109/ICCMC51019.2021.9418375, 2021.

[6] S. Jain and D. Ramesh, "Machine Learning convergence for weather-based crop selection," IEEE International Students' Conference on Electrical, Electronics and Computer Science(SCEECS), pp. 1-6, doi: 10.1109/SCEECS48394.2020.75, 2020.

[7] Suresh, P. Ganesh Kumar and M. Ramalatha, "Prediction of major crop yields of Tamilnadu using K-means and Modified KNN," 3rd International Conference on Communication and Electronics Systems (ICCES), pp. 88-93, doi: 10.1109/CESYS.2018.8723956, 2018.

[8] S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika and J. Nisha, "Crop recommendation system for precision agriculture," Eighth International Conference on Advanced Computing (ICoAC), pp. 32-36, doi: 10.1109/ICoAC.2017.7951740, 2017.

[9] R. Kumar, M. P. Singh, P. Kumar and J. P. Singh, "Crop Selection Method to maximize crop yield rate using machine learning technique," International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), pp. 138- 145, doi: 10.1109/ICSTM.2015.7225403, 2015.

[10] S. M. PANDE, P. K. RAMESH, A. ANMOL, B. R. AISHWARYA, K. ROHILLA and K. SHAURYA, "Crop Recommender System Using Machine Learning Approach," 5th International Conference on Computing Methodologies and Communication (ICCMC).

[11] Raj, Angu & Balashanmugam, Dr.Thiyaneswaran & Jayanthi, J & Yoganathan, N & Srinivasan, PCrop. . Recommendation on Analyzing Soil Using Machine Learning. Turkish Journal of Computer and Mathematics Education (TURCOMAT). 12. 1784-1791, 2021.

[12] T. Ragunthar, S. Selvakumar and G. Ilamurugan, " Counsel System for Effective Farming using Data Mining Algorithm", International Journal of Pure And Applied Mathematics, Volume 117 No. 21, 921-924 ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (online version), 2017.

[13] Shridhar Mhaiskar, Chinmay Patil, Piyush Wadhai, Aniket Patil and Vaishali Deshmukh, “A Survey on Predicting Suitable Crops for Cultivation Using IoT ”, International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization), Website: www.ijircce.com, Vol. 5, Issue 1, January 2017.

[14] Lakshmi N, Priya M, Mrs Sahana Shetty and Mr. Manjunath C.R, "Crop Recommendation System for Precision Agriculture ", International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653; IC Value: 45.98; SJImpact Factor: 6.887, Volume 6 Issue V, May 2018.

[15] Lakshmi. N, Priya. M, Mrs. Sahana Shetty and Mr. Manjunath C.R, "Crop Recommendation System for Precision Agriculture ", International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653; IC Value: 45.98; SJImpact Factor: 6.887, Volume 6 Issue V, May 2018.

[16] M.V.R. Vivek, D.V.V.S.S. Sri Harsha and P. Sardar Maran, “A Survey on Crop Recommendation Using Machine Learning”, International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277- 3878, Volume-7, February 2019.

[17] Patel, Harsh & Prajapati, Purvi. “Study and Analysis of Decision Tree Based Classification Algorithms”. International Journal of Computer Sciences and Engineering. 6. 74-78. 10.26438/ijcse/v6i10.7478, 2018.

[18] Kaviani, Pouria & Dhotre, Sunita. “Short Survey on Naive Bayes Algorithm”. International Journal of Advance Research in Computer Science and Management. 04, 2017.

[19] Evgeniou, Theodoros & Pontil, Massimiliano. Support Vector Machines: Theory and Applications. 2049. 249-257. 10.1007/3-540-44673-7_12, 2001.

[20] Rong, Shen & Bao-wen, Zhang. “The research of regression model in machine learning field”. MATEC Web of Conferences. 176. 01033. 10.1051/matecconf/201817601033, 2018.

[21] Schonlau, Matthias, and Rosie Yuyan Zou. “The Random Forest Algorithm for Statistical Learning.” The Stata Journal, vol. 20, no. 1, pp. 3–29, doi:10.1177/1536867X20909688, 2018.

[22] Santhanam, Ramraj & Uzir, Nishant & Raman, Sunil & Banerjee, Shatadeep. Experimenting XGBoost Algorithm for Prediction and Classification of Different Datasets, 2017.

[23] Weather Forecast API https://openweathermap.org/api

[24] https://www.kaggle.com/atharvaingle/crop-recommendation-dataset

[25] https://www.kaggle.com/abhinand05/crop-production-in-india