SUSTAINABLE FERTILIZER USAGE OPTIMIZER FOR HIGHER YIELD

Mr. V.Murugan[1] , Ujvala.M[2], RAVI.P[3], Noorein Fatima[4], Nithin.N[5]

[1]*Assistant Professor, Department of IT, TKR College of Engineering and Technology, Telangana, India*
[2,3,4,5]*B.Tech Students, Department of IT, TKR College of Engineering and Technology, Telangana, India*

---***---

Abstract - *Agricultural productivity is highly influenced by soil nutrients, climatic conditions, and farming practices such as fertilizer usage and crop rotation. However, many farmers still rely on traditional knowledge and intuition for crop selection and fertilizer application, which often leads to reduced yield, higher input costs, and long-term soil degradation. To address this issue, this paper proposes a Sustainable Fertilizer Usage Optimizer for Higher Yield, an intelligent web-based agriculture recommendation system that uses machine learning and ensemble learning techniques. The system predicts soil characteristics, recommends the most suitable crop, estimates expected yield, and suggests the optimal fertilizer type and quantity based on soil parameters such as Nitrogen (N), Phosphorus (P), Potassium (K), pH, moisture, and weather conditions including temperature and rainfall. Advanced models such as LightGBM, XGBoost, AdaBoost, Extra Trees, and Gradient Boosting are utilized, and their performance is enhanced using ensemble strategies like bagging, boosting, and stacking. Cross-validation and hyperparameter tuning are applied to improve accuracy and reduce overfitting. The proposed system supports sustainable farming by minimizing excessive fertilizer usage, improving crop yield, and enabling farmers to make data-driven decisions. This approach enhances productivity, reduces environmental impact, and increases long-term profitability for farmers.*

KEYWORDS:PrecisionAgriculture,Crop ecommendation, Fertilizer Optimization, Yield Prediction, Soil Nutrient Analysis

1. INTRODUCTION

Agriculture plays a vital role in ensuring food security and supporting the economy of developing countries such as India. A major portion of the population depends on farming as their primary source of income. However, agricultural productivity is influenced by multiple factors such as soil nutrient levels, climatic variations, irrigation availability, crop rotation, and fertilizer management. In many cases, farmers rely on traditional knowledge and experience for selecting crops and applying fertilizers. Although these practices are useful, they often fail to provide accurate decisions under changing environmental conditions and lead to low yield, increased production cost, and soil fertility degradation.

Recent advancements in precision agriculture and data-driven farming have enabled the use of machine learning (ML) techniques to improve crop productivity. ML models can analyze soil parameters such as Nitrogen (N), Phosphorus (P), Potassium (K), pH, moisture, and climatic conditions such as rainfall and temperature to recommend suitable crops, predict yield, and suggest fertilizer requirements [1]. Studies show that ensemble learning methods such as Gradient Boosting, Random Forest, and XGBoost provide better performance compared to traditional ML models due to their ability to handle nonlinear agricultural datasets and improve prediction accuracy [2].

Fertilizer usage is another major concern in modern agriculture. Excessive and improper fertilizer application results in nutrient imbalance, groundwater pollution, reduced soil quality, and environmental damage. Sustainable fertilizer recommendation systems can help in reducing these issues by providing optimal fertilizer type and quantity based on soil nutrient deficiency and crop requirement [3]. Similarly, yield prediction plays a crucial role in planning harvesting strategies, supply chain management, and market decision-making for farmers. Yield prediction using historical yield records, soil health parameters, and weather conditions has been widely researched and is considered an effective approach for improving profitability [4].

This project proposes a Sustainable Fertilizer Usage Optimizer for Higher Yield, a web-based intelligent recommendation system that integrates multiple machine learning algorithms such as LightGBM, XGBoost, AdaBoost, Extra Trees, and Gradient Boosting. The system supports four major functions: Soil Prediction, Crop Recommendation, Fertilizer Suggestion, and Yield Prediction. Ensemble learning techniques like bagging, boosting, and stacking are applied along with cross-validation and hyperparameter tuning to enhance model accuracy and generalization. The proposed system provides farmers with reliable, real-time recommendations to increase yield, reduce input cost, and promote sustainable agricultural practices.

1.1 Motivation

The motivation behind this work is to support farmers by providing an intelligent decision-making system for crop selection and fertilizer optimization. Farmers often face uncertainty due to unpredictable climate, soil nutrient imbalance, and lack of scientific recommendations. By integrating soil and weather parameters with machine learning models, this system aims to improve productivity, reduce fertilizer misuse, and enhance sustainability.

1.2 Problem Statement

Farmers often make crop and fertilizer decisions based on intuition or traditional methods rather than data-driven analysis. This results in low yield, increased cultivation cost, excessive fertilizer usage, and long-term damage to soil quality. Therefore, there is a need for an intelligent system that can:

Recommend the best crop based on soil conditions and climate. Predict expected yield using historical and real-time parameters. Suggest optimal fertilizer type and quantity with sustainability.

2. PROPOSED SYSTEM

Pproposed Hybrid Multi-Modal Deep Learning The proposed system is a web-based intelligent agriculture recommendation platform designed to assist farmers in selecting suitable crops, predicting yield, and optimizing fertilizer usage in a sustainable manner. The system utilizes soil parameters such as Nitrogen (N), Phosphorus (P), Potassium (K), and pH along with climatic parameters such as temperature, rainfall, and humidity to generate accurate predictions. Unlike traditional approaches, this system integrates multiple machine learning models and ensemble learning techniques to improve reliability and prediction accuracy. The core objective is to provide a complete decision-support system that reduces excessive fertilizer usage, improves crop yield, and enhances long-term soil health. The proposed platform includes four main functional modules: Soil Prediction, Crop Recommendation, Fertilizer Suggestion, and Yield Prediction.

2.1 Soil Prediction Module

The soil prediction module is responsible for analysing soil fertility and classifying the soil condition based on essential nutrient parameters. The system takes soil input values such as Nitrogen, Phosphorus, Potassium, and pH level, and processes them using tree-based machine learning classifiers. These models learn patterns from historical soil datasets and predict the soil category or fertility level. This module plays a foundational role because accurate soil classification directly improves the performance of crop recommendation and fertilizer optimization. The output of this module provides a clear understanding of soil nutrient deficiency and soil suitability for different crops.

2.2 Crop Recommendation Module

The crop recommendation module predicts the most suitable crop for cultivation based on the current soil condition and weather environment. It uses advanced machine learning algorithms such as LightGBM, XGBoost, AdaBoost, Extra Trees, and Gradient Boosting to generate crop recommendations. Ensemble learning strategies such as bagging, boosting, and stacking are applied to enhance the model's accuracy and stability. The system ensures better generalization by applying cross-validation and hyperparameter tuning during training. This module helps farmers choose crops that have higher survival probability, better yield potential, and maximum profitability under the given environmental conditions.

2.3 Fertilizer Suggestion Module

The fertilizer suggestion module recommends the most suitable fertilizer type and quantity required for the selected crop based on soil nutrient deficiency. This module focuses on sustainable farming by preventing overuse of fertilizers and maintaining soil nutrient balance. The prediction model compares the current NPK levels with the ideal nutrient requirement for the recommended crop and identifies which nutrient is deficient or excessive. Based on this analysis, the system suggests appropriate fertilizer inputs that can restore soil fertility without harming the environment. This approach reduces input cost for farmers and also minimizes soil degradation and groundwater pollution caused by improper fertilizer application.

2.4 Yield Prediction Module

The yield prediction module estimates the expected crop yield based on soil nutrients, weather parameters, and historical yield records. This module is implemented using regression-based ensemble learning algorithms that can handle nonlinear relationships between agricultural features and yield output. By analysing patterns from past yield datasets, the system predicts yield in terms of quintals per hectare. Yield prediction is useful for farmers to plan harvesting strategies, manage resources effectively, and estimate future profit margins. The module also supports scalability and can be improved further by integrating real-time weather updates and market datasets for economic planning.

3. IMPLEMENTATION DETAILS

The implementation of the proposed Sustainable Fertilizer Usage Optimizer for Higher Yield is carried out as a modular machine learning based web application. The system is designed to collect agricultural datasets, preprocess and normalize the data, train multiple machine learning models, and generate predictions through an interactive user interface. The complete implementation is divided into four major stages: data collection, data preprocessing, model training and evaluation, and system deployment. Each stage is carefully designed to ensure accuracy, scalability, and usability for farmers.

3.1 Data Collection

The first stage of implementation involves collecting agricultural datasets from reliable sources. The collected data includes soil nutrient parameters such as Nitrogen, Phosphorus, Potassium, and pH, along with environmental parameters such as temperature, humidity, and rainfall. In addition, historical crop yield data is also included to support yield prediction. The data is collected from multiple formats such as CSV files, spreadsheets, and public agricultural repositories. The collected datasets are then stored and organized to ensure compatibility with machine learning workflows. This stage is important because the performance of the system depends heavily on the quality and diversity of the dataset used.

3.2 Data Preprocessing

After collecting the dataset, preprocessing is performed to clean and prepare the data for machine learning training. This stage includes removing missing values, handling duplicate records, correcting inconsistent values, and formatting data into a structured form. Feature scaling and normalization techniques are applied where required, especially for algorithms that are sensitive to feature ranges. Unwanted columns are removed, and only relevant attributes such as NPK, pH, temperature, rainfall, and humidity are retained. The dataset is then split into training and testing sets. This preprocessing step ensures that the model learns meaningful patterns and improves overall prediction accuracy.

3.3 Model Training and Evaluation

In this stage, multiple machine learning models are trained for each prediction module. For crop recommendation and soil prediction, classification models such as LightGBM, XGBoost, AdaBoost, Extra Trees, and Gradient Boosting are trained using labeled datasets. For yield prediction, regression-based models are trained to estimate yield output. Ensemble learning techniques such as bagging, boosting, and stacking are applied to enhance robustness and reduce variance. Cross-validation is used during training to prevent overfitting and ensure that the models perform well on unseen data. Hyperparameter tuning is performed to select the best model configuration. The trained models are evaluated using accuracy, precision, recall, and F1-score for classification tasks, and MAE, RMSE, and R^2 score for regression tasks.

3.4 System Deployment and Architecture Integration

After model training, the best-performing models are integrated into a web-based platform. The system allows farmers to enter soil nutrient values and weather parameters through a user-friendly interface. Once the input is provided, the backend loads the trained models and generates outputs such as soil type prediction, recommended crop, fertilizer suggestion, and expected yield. The results are displayed instantly to the user in a clear format. The proposed system architecture is integrated at this stage to show the flow of data from user input to prediction output. This architecture diagram represents the interaction between the user interface, preprocessing layer, machine learning models, ensemble module, and the final recommendation output.

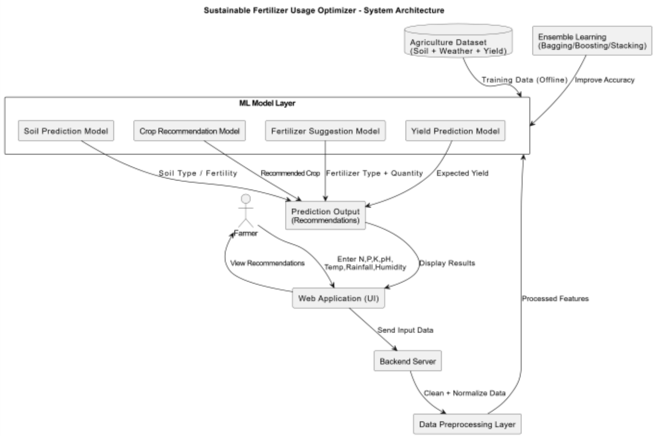

Fig -1: Proposed System Architecture

4. RESULTS AND PERFORMANCE ANALYSIS

This section discusses the performance of the proposed Sustainable Fertilizer Usage Optimizer for Higher Yield system. The system was evaluated based on its ability to predict soil type, recommend the most suitable crop, suggest fertilizer requirements, and estimate crop yield. Multiple machine learning models such as LightGBM, XGBoost, AdaBoost, Extra Trees, and Gradient Boosting were trained and tested using agricultural datasets containing soil nutrient parameters and weather conditions. The models were evaluated using standard classification and regression metrics to ensure reliability and generalization.

For soil prediction and crop recommendation, classification models were evaluated using accuracy, precision, recall, and F1-score. The results show that ensemble-based algorithms performed better compared to single traditional models due to their ability to handle nonlinear relationships and feature interactions. Among the tested models, LightGBM and XGBoost achieved higher accuracy and stable performance across cross-validation folds. The use of bagging and boosting helped reduce variance and improve overall robustness. The stacking approach further enhanced predictive performance by combining the outputs of multiple base learners.

For yield prediction, regression-based ensemble models were evaluated using Mean Absolute Error (MAE), Root

Mean Square Error (RMSE), and R^2 score. The results indicate that Gradient Boosting and Extra Trees regression models produced lower error values and better yield estimation accuracy. This confirms that ensemble learning techniques are effective for yield prediction tasks where the relationship between soil, climate, and yield is complex and highly nonlinear.

The system outputs are displayed through a user-friendly web interface. When the farmer enters soil nutrient values and climatic parameters, the system generates predictions such as soil type (example: Loamy), recommended crop, fertilizer suggestion, and expected yield. The result page demonstrates that the proposed platform provides accurate and real-time recommendations, making it practical for real-world usage.

The overall experimental results validate that the proposed system can support farmers in making data-driven agricultural decisions. By combining soil and climate analysis with advanced ensemble learning methods, the system improves prediction accuracy, reduces fertilizer misuse, and increases productivity. The integration of multiple models and tuning techniques ensures consistent performance and scalability for different regions and crops.

Fig -2: Result Page Showing Predicted Soil Type (Loamy)

5. CONCLUSION

This paper presented a Sustainable Fertilizer Usage Optimizer for Higher Yield, an intelligent web-based agriculture recommendation system designed to support farmers in making accurate and sustainable farming decisions. The proposed system integrates advanced machine learning models such as LightGBM, XGBoost, AdaBoost, Extra Trees, and Gradient Boosting to perform soil prediction, crop recommendation, fertilizer suggestion, and yield prediction. By applying ensemble learning techniques including bagging, boosting, and stacking, the system achieves improved accuracy, robustness, and generalization compared to traditional approaches.

The implementation demonstrates that the system can effectively analyze soil nutrient parameters and climatic conditions to provide real-time recommendations. The fertilizer optimization module helps reduce excessive fertilizer usage, thereby lowering cultivation cost and minimizing environmental impact. The yield prediction module further supports farmers by estimating expected production, enabling better planning and profitability analysis. Overall, the proposed system contributes to precision agriculture by enhancing productivity, maintaining soil health, and promoting sustainable farming practices.

6. FUTURE WORK

The proposed system can be further enhanced in several ways to improve its real-world applicability and scalability. In future, real-time data collection can be integrated using IoT sensors to automatically capture soil moisture, temperature, and nutrient levels without manual input. The system can also be extended by incorporating satellite imagery and remote sensing data for large-scale farm monitoring and crop health assessment. Additionally, market price prediction and demand forecasting modules can be included to help farmers make more profitable decisions based on both yield and economic conditions. Support for regional languages and a mobile application version can increase accessibility for rural farmers. Finally, the system can be trained with larger region-specific datasets to improve accuracy across different soil types, crops, and climatic zones.

REFERENCES

[1]Ayesha Khaliq, Atif Khan, Salman Jan, Muhammad Umair, Asad Gulshair, Ahmad Ali, Usman Ali ShahAI-Driven Smart Agriculture: An Integrated Approach for Soil Analysis, Irrigation, and Crop-Fertilizer Recommendations Digital Object Identifier 10.1109/ACCESS.2024.Doi Number

[2] R. Gandhi, S. Nimbalkar, N. Yelamanchili, and S. Ponkshe, "Crop recommendation system using machine learning," in Proc. IEEE Int. Conf. on Computing, Communication and Automation (ICCCA), 2016, pp. 1–5.

[3] M. M. Rahman, M. S. Islam, M. A. Rahman, and M. S. Hossain, "A machine learning approach for crop yield prediction using climatic and soil parameters," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 11, no. 9, pp. 1–7, 2020.

[4] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," Computers and Electronics in Agriculture, vol. 147, pp. 70–90, 2018.

[5] S. Jeong, S. M. Jang, and S. Kim, "A crop yield prediction method based on machine learning and climate data," Applied Sciences, vol. 10, no. 21, pp. 1–15, 2020.

[6] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, 2016, pp. 785–794.

[7] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T. Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in Proc. 31st Int. Conf. on Neural Information Processing Systems (NeurIPS), 2017, pp. 3146–3154.

[8] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[9] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Annals of Statistics, vol. 29, no. 5, pp. 1189–1232, 2001.

[10] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," Machine Learning, vol. 63, no. 1, pp. 3–42, 2006.