A Data-Driven Framework for Crop and Fertilizer Recommendation in Smart Agriculture

[1]D.Charan, [2]B.Sri Umamaheshwar, [3]M.T.S.Ganesh, [4]K.H.S.Krishna, [5] Dr.R.J.Mathew
[1234]B.Tech Student, [5]Associate Professor
in Dept. of CSE at Srinivasa Institute Of Engineering And Technology, (Autonomous), NH-216, Cheyyeru(V), Amalapuram -533216.

ABSTRACT:

This research introduces a smart farming solution to aid farmers in selecting the most appropriate crops based on soil conditions and seasonal variations. Utilizing machine learning algorithms such as Random Forest and Decision Tree, the system analyzes soil parameters like nitrogen, phosphorus, potassium, pH, humidity, and rainfall to recommend optimal crops for cultivation. The model also supports fertilizer recommendations and detects plant diseases using image processing techniques. By integrating data mining with agricultural intelligence, this system empowers farmers to make data-driven decisions that enhance productivity and reduce crop failure risks. It ultimately aims to modernize traditional agricultural practices with a reliable, web-based application, improving crop yield and supporting sustainable farming.

Keyword: Soil Nutrient Analysis, Machine Learning, Smart Farming

INTRODUCTION

Agriculture remains the backbone of many economies, yet modern challenges like unpredictable weather, poor crop selection, and soil degradation hinder productivity. Traditional farming methods, often based on intuition or past practices, lack the precision needed for optimal yield. This project introduces a data-driven solution that leverages machine learning to recommend suitable crops based on soil nutrient data and seasonal patterns. By analyzing critical soil parameters such as nitrogen, phosphorus, potassium, and pH along with environmental conditions like humidity and rainfall, the system offers accurate crop suggestions. Additionally, it provides fertilizer guidance and identifies plant diseases through image classification techniques. The use of algorithms such as Random Forest and Decision Tree enhances prediction accuracy. This intelligent platform serves as a digital assistant for farmers, empowering them to make informed decisions, reduce resource wastage, and boost agricultural efficiency. It is a step forward in promoting precision agriculture and sustainable crop management.

RELATED WORK

Several researchers have explored machine learning techniques for agricultural optimization. Pudumalar et al. (2016) developed a crop prediction model using ensemble methods including Random Trees, KNN, CHAID, and Naïve Bayes. Their majority voting system enhanced accuracy to 88%, demonstrating the value of combining algorithms to improve predictive power. This approach minimized the risk of single-model errors, making it robust for real-world applications.

Gandge and Sandhya (2017) presented a comparative study of various machine learning algorithms such as Multiple Linear Regression (90–95% accuracy for rice), Decision Tree (ID3 for soybean), and SVM for general crop prediction. Their findings highlight that although SVM and neural networks offer high accuracy, traditional models like regression remain competitive depending on data structure.

Rakesh Kumar et al. (2018) proposed a hybrid model integrating ANN, SVM, Random Forest, and GBDT for crop selection based on yield data. Their work emphasized the importance of time-based crop rotation and yield rate analysis for optimized agricultural planning.

Rakesh Shirsath et al. (2019) introduced a subscription-based crop recommendation platform that maintained farmer-specific data and past crop records. Using artificial neural networks, it tailored crop suggestions and included a feedback system to adapt future recommendations based on user responses.

Jinchuan Zhao and Jian-Xin Guo (2020) worked on an agricultural decision system using Big Data analytics. Their model utilized modules like a deduction engine, data builders, and a knowledge base to extract actionable insights from massive agricultural datasets.

TABLE1. Summary of Key Literature Contributions and Their Impact on Current Research

Author(s)	**Contribution**	**Impact on Research**
S. Pudumalar et al. (2016)	Developed an ensemble model combining Random Trees, KNN, CHAID, and Naïve Bayes using majority voting for crop prediction.	Improved prediction accuracy (88%) by reducing the effect of single-model misclassification.
Yogesh Gandge & Sandhya (2017)	Conducted a comparative study on ML algorithms like SVM, Regression, and ID3 for predicting crop yield and disease.	Highlighted algorithm-specific strengths; emphasized the need for dataset-specific algorithm selection.
Rakesh Kumar et al. (2018)	Proposed use of ANN, SVM, Random Forest, GBDT, and Regularized Gradient Forest for crop yield estimation and season-specific planting.	Showcased hybrid modeling as a path to improve yield prediction and seasonal planning.
Rakesh Shirsath et al. (2019)	Created a subscription-based recommendation system with personalized user data, feedback loops, and crop history tracking using neural networks.	Promoted adaptive learning and farmer-specific recommendations through feedback integration.
Jinchuan Zhao & Jian-Xin Guo (2020)	Designed a Big Data-based agricultural decision system using modules like inference engines, data constructors, and knowledge bases.	Demonstrated how large-scale data and domain-specific inference improve decision-making accuracy in agricultural systems.

PROPOSED APPROACH

The proposed system offers a machine learning-based crop recommendation solution tailored to seasonal and soil conditions. It begins with farmers collecting soil samples from their fields and submitting them for laboratory analysis. The system uses the soil nutrient values—Nitrogen (N), Phosphorus (P), Potassium (K), pH level, temperature, humidity, and rainfall—as key input features. These parameters are essential in determining the suitability of crops for a given region.

The data collected over years from reliable agricultural sources (e.g., Kaggle Soil Health datasets) is used to train classification models. The Random Forest and Decision Tree algorithms are employed to provide recommendations with high accuracy and computational efficiency. Based on the input, the system predicts suitable crops that can yield maximum productivity under the current conditions.

In addition to crop suggestions, the system offers fertilizer recommendations based on nutrient deficiencies detected in the soil. An integrated module also allows plant disease identification through image processing using a deep learning-based CNN model. Once a disease is detected, the system suggests remedies to treat it organically or chemically.

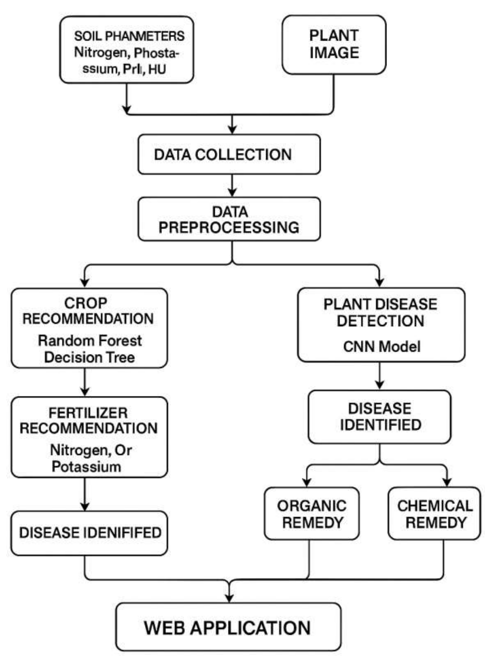

Figure 1: Crop recommendation system

METHODOLOGIES

Data Collection and Preprocessing

The system uses historical datasets obtained from sources like Kaggle and agricultural departments. The primary dataset includes features such as Nitrogen (N), Phosphorus (P), Potassium (K), pH value, humidity, rainfall, and temperature. These features are cleaned and normalized to ensure consistency and reduce bias.

2. Crop Recommendation Module

The core of this module is based on Random Forest and Decision Tree algorithms. These supervised learning models are trained using the soil nutrient and climate data to classify the most suitable crops. Random Forest is chosen for its ability to reduce overfitting and handle large feature sets, while Decision Tree offers clear interpretability.

3. Fertilizer Recommendation System

This component analyzes the gap between optimal and current nutrient levels of the soil. It compares real-time soil data with crop-specific fertilizer requirements and recommends either Nitrogen-rich, Phosphorus-rich, or Potassium-rich fertilizers accordingly.

4. Soil Classification

Using the same feature set, soil types are classified to match appropriate crops. This step refines the prediction by ensuring crop compatibility with soil texture and nutrient profile.

5. Crop Disease Detection

The disease module uses a CNN-based deep learning model (ResNet9) trained on thousands of plant leaf images. Farmers upload an image, and the system classifies the disease and recommends suitable organic or chemical treatment. Image preprocessing includes resizing, normalization, and tensor conversion using PyTorch.

6. Web Application Deployment

Flask is used to build the user-friendly web interface. It allows farmers to input data, receive recommendations, and visualize results dynamically. Backend models are integrated with the front-end via API endpoints.

RESULTS

The proposed system was tested using real-world agricultural datasets and image samples to validate its performance across crop recommendation, fertilizer suggestion, and plant disease detection. The crop prediction module, trained using Random Forest and Decision Tree algorithms, achieved high classification accuracy. Random Forest consistently outperformed other models with over 95% accuracy in recommending crops suitable for specific soil and environmental conditions.

The fertilizer recommendation system effectively analyzed discrepancies between existing soil nutrient levels and optimal crop requirements. Based on this analysis, it accurately advised farmers on the type of fertilizer (Nitrogen-rich, Phosphorus-rich, or Potassium-rich) needed to balance the soil composition, thereby increasing productivity.

For disease detection, the system employed a convolutional neural network (ResNet9), which correctly identified plant diseases from uploaded leaf images with significant precision. The model was capable of distinguishing between similar visual symptoms and suggesting organic or chemical treatments based on the diagnosis.

DISCUSSION

The integrated system presented in this project addresses several key challenges faced by modern agriculture, such as inappropriate crop selection, nutrient imbalance, and delayed disease detection. By leveraging machine learning models and data-driven approaches, the system enhances agricultural decision-making with minimal human intervention. The results obtained during testing confirm that Random Forest performs better than Decision Tree due to its ensemble nature and ability to handle noisy or missing data. Its accuracy and robustness make it ideal for practical implementation in rural settings.

One of the major strengths of the system is its adaptability. It can recommend crops not only based on soil nutrients but also by factoring in real-time weather conditions like temperature and humidity. Furthermore, the fertilizer module dynamically adjusts to nutrient deficiencies, providing personalized guidance for improved yield. The disease detection module, powered by deep learning, adds another layer of functionality, reducing dependency on expert diagnosis and minimizing losses due to unidentified infections.

However, challenges remain, such as ensuring continuous dataset updates to reflect seasonal variations and regional differences. Additionally, internet accessibility and user training in rural areas could affect system adoption. Future enhancements like multilingual support and mobile app integration could further expand the platform's reach and usability.

CONCLUSION

This work successfully demonstrates a smart, integrated solution for crop recommendation, fertilizer guidance, and plant disease detection using advanced machine learning and deep learning techniques. By analyzing soil nutrient data, environmental conditions, and visual plant symptoms, the system provides reliable suggestions to farmers, enabling informed agricultural decisions. The use of Random Forest and Decision Tree algorithms ensures high accuracy in crop prediction, while CNN-based disease identification allows timely interventions to reduce crop damage.

The system not only aids in maximizing productivity but also promotes sustainable farming by optimizing resource usage. Its user-friendly web interface ensures ease of access even for users with limited technical knowledge. While the results are promising, real-world deployment may require additional features such as mobile support and multilingual interfaces to enhance usability in rural areas.

REFERENCES

1. Pudumalar, S., Ramya, R., Sankar, R., Vignesh, P., Gokulakrishnan, B. and Harini, R., 2016. Crop recommendation system for precision agriculture. *Procedia Computer Science*, 93, pp.931–936.
2. Gandge, Y. and Sandhya, A., 2017. A review of data mining techniques for crop yield prediction. *International Journal of Computer Science and Information Technologies*, 8(2), pp.256–258.
3. Kumar, R., Singh, M.P., Kumar, P. and Singh, J.P., 2018. Crop selection method to maximize crop yield rate using machine learning technique. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 3(2), pp.12–18.
4. Shirsath, R. and Kamble, D., 2019. A smart farming: Crop recommendation system using machine learning. *International Journal of Engineering Research and Technology*, 8(11), pp.1–5.
5. Zhao, J. and Guo, J.X., 2020. Application of big data analysis technology in agricultural intelligent decision-making system. *IOP Conference Series: Earth and Environmental Science*, 558(1), p.012112.
6. Singh, V., Rathore, R. and Rajpurohit, V.S., 2017. Machine learning approach for crop yield prediction. *International Journal of Advanced Research in Computer Science*, 8(5), pp.1564–1567.
7. Chlingaryan, A., Sukkarieh, S. and Whelan, B., 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture*, 151, pp.61–69.
8. Patel, K. and Dholakia, M., 2020. Smart crop recommendation system using machine learning. *International Journal of Scientific & Technology Research*, 9(2), pp.5221–5223.
9. Pant, R., Sahoo, P.K. and Saha, P.K., 2019. Fertilizer recommendation system for precision agriculture using machine learning. *Journal of Ambient Intelligence and Humanized Computing*, 10(6), pp.2375–2384.
10. Kamilaris, A. and Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, pp.70–90.
11. Brahimi, M., Boukhalfa, K. and Moussaoui, A., 2017. Deep learning for tomato diseases: Classification and symptoms visualization. *Applied Artificial Intelligence*, 31(4), pp.299–315.
12. Mohanty, S.P., Hughes, D.P. and Salathé, M., 2016. Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, p.1419.
13. Kohli, A., Joshi, A. and Parmar, B., 2021. Intelligent fertilizer recommendation system using ensemble learning. *Procedia Computer Science*, 185, pp.191–198.
14. Srivastava, A., 2020. Crop prediction using machine learning algorithms. *International Journal of Engineering Research & Technology*, 9(4), pp.87–91.
15. Sharma, P., Jain, A. and Singh, A., 2020. Predictive analysis for crop recommendation using machine learning techniques. *International Journal of Scientific Research in Computer Science and Engineering*, 8(1), pp.30–35.
16. Ghosh, R., 2019. Deep learning-based plant disease detection using image processing. *International Journal of Innovative Technology and Exploring Engineering*, 8(11), pp.223–227.
17. Chauhan, A. and Mathur, D., 2021. Weather-based smart crop prediction using machine

learning. *Journal of Emerging Technologies and Innovative Research*, 8(6), pp.293–297.
18. Ahmed, R., 2021. A survey on crop recommendation using machine learning and IoT. *Materials Today: Proceedings*, 47(4), pp.7303–7307.
19. Jain, R. and Taneja, M., 2020. Design and implementation of web-based agricultural recommendation system. *International Journal of Computer Applications*, 177(22), pp.8–12.
20. Rajalakshmi, P. and Arunkumar, P., 2017. Intelligent crop recommendation system using machine learning techniques. *International Journal of Advanced Research in Computer and Communication Engineering*, 6(10), pp.529–534.