Crop Recommendation Based on Soil and Weather Data

Sudhaharniie Thevar[1], Armaan Mujawar[2], Sohail Shaikh[3], Prini Rastogi[4]

The Department of CTIS,
Ajeenkya DY Patil University, Pune, India.

Abstract

Among the factors that have caused agriculture to have low productivity are climate variability, imbalance of nutrients in soils, and poor choice of crops, this paper presents a Smart Crop Advisory System that employs machine learning with real-time environmental information to make agricultural decisions. The system analyses such parameters as nitrogen (N), phosphorus (P), potassium (K), pH, temperature, humidity, and rainfall and recommends appropriate crops based on a Random Forest machine learning model. The OpenWeatherMap API provides weather data in real-time to improve the quality of forecasts and provide an understanding of the climate. The system also has the capability of yield prediction, irrigation instructions and soil analysis features, which are an interactive Flask-based dashboard. Experimental testing of this system has achieved an accuracy of approximately 94% confirming the effectiveness of the system in precision agriculture.

Keywords: Precision agriculture, Machine learning, Crop recommendation, Smart farming, Random Forest.

1. Introduction

There are several systems on the market with limited implementation and scope even though smart agricultural technologies are developing fast. Many studies have explored mainly single functionalities like crop prediction with machine learning model, soil nutrient analysis, or environmental monitoring with the help of IoT devices [1], [4]. Even though these strategies have been performing well under controlled environments, the strategies normally lack real-time data, which is critical in controlling dynamical agricultural settings, as affected by climate change. In addition, other systems related to the IoT may constantly monitor the environmental conditions and parameters, such as soil moisture, temperature, and humidity, but they do not tend to provide predictive analytics and smart decision-support systems that can further transform raw data into farm-indicative information [9].

The other major constraint that has been noted in the current systems is that they fail in providing user-friendliness coupled with interactive visualization tools, thus limiting their use particularly among small and marginal farmers. There are many proposed machine learning models that are still only prototypes and have not provided a solution that can integrate data analysis, prediction, and visualization into a user-friendly interface. Also, the lack of interconnection between the machine learning model and other external services, such as weather forecasts and market prices, makes it not as useful for real-world decision-making.

To overcome the issues, the given project proposes a complex system called the Smart Crop Advisory System, in which machine learning techniques are embedded along with real-time data providers in the form of a single system. The system utilizes the results of the machine learning model to provide precise suggestions for crops based on the soil parameters such as temperature, humidity, rainfall. In addition to this, real-time weather and market information are also provided to the system for better decision-making. The system has an interactive and user-friendly interface that provides the user with essential information such as the suitability of the crops, soil analysis, and yield prediction.

The offered system will fill the gap between theoretical research and practical solutions of agriculture based on combining predictive analytics and real-time environmental data, as well as visualization. This integrated strategy is aligned with the recent studies on the importance of adaptive, data-intensive, and user-friendly decision support systems to increase agricultural productivity, use resources effectively, and survive the risks of unpredictable weather patterns

2. Literature Review

In recent literature in the smart agriculture sector, machine learning approaches have been examined when it comes to enhancing crop productivity and decision-making. Random Forest, Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Deep Learning models are algorithms applied by different researchers to complete various tasks by classifying crops, predicting their yield, identifying diseases, and analyzing soil, among others [1]. The most popular among them has been ensemble learning techniques as they are more accurate and robust.

In a systematic literature review of machine learning-based predictions on crop yields, T. van Klomkenburg et al. concluded that ensemble models, in particular, Random Forest and Gradient Boosting, outperformed the traditional regression models in the analysis of more complex agricultural data [1]. A nonlinear interaction of environmental and soil parameters can be described with the help of such models that will contribute to the improved results of the predictions.

At the same time, the Internet of Things (IoT) technologies have enabled technological progress in agricultural monitoring systems. Akhter and Sofi proposed an intelligent agriculture system which is IoT-based and enables the continual monitoring of the soil moisture levels, temperature, and humidity levels in the environment using wireless sensor networks [9]. These kinds of systems have the advantage of real time data capture, something that is very vital in precision farming practices. Internet of Things technologies are also important in smart agriculture because they allow devices and sensors in the agricultural environment to communicate with monitoring systems in real time [9], [12].

Recommendation systems of crops based on machine learning model trained on soil nutrient value and climatic conditions are also worked on in other studies. Such systems gauge the levels of such parameters as nitrogen (N), phosphorus (P), potassium (K), temperature, humidity, rainfall and PH to recommend the best crops to be planted in certain regions [3], [4]. A combination of these parameters has been proved to increase the accuracy of crop choice and agricultural planning significantly.

Furthermore, accuracy in agricultural research studies are indicative of the strength of artificial intelligence as far as smart decision-making is concerned. AI-based systems can be used to analyze various environmental and soil data sets to assist in implementing effective crop management practices and optimization of resources [2]. Moreover, the use of predictive models coupled with weather forecasting information has been shown to improve the accuracy of crop prediction systems [5]. Deep learning techniques have also been applied for crop disease detection and classification, improving early identification and management of plant diseases [6], [8].

The smart agriculture solutions currently in place have been successful despite these advances, however, a number of constraints remain. Most researchers concentrate on particular aspects like crop prediction models or IoT-based monitoring systems without a single platform [11]. Also, insufficient user-friendly interfaces and feasible advisory systems reduce the availability of such technologies to farmers, especially when they have minimal knowledge of technology. [11]

3. Proposed System

The Smart Crop Advisory System is a web-based smart farming system that will support farmers in enhancing their crop yield and make data-driven agricultural decisions. The conventional form of farming is usually experience and information based and hence may result in poor choice of crop and poor utilization of resources [11]. To overcome this shortcoming, the system combines machine learning model with real-time environmental data to give realistic and effective suggestions via a single digital interface [5].

The system enables farmers to input some of the key soil and environmental variables into the system with ease through a user-friendly dashboard including nitrogen (N), phosphorus (P), potassium (K), soil pH, temperature, humidity, and rainfall. These parameters are then fed into the trained machine learning model which is used to recommend the most appropriate crops, based on the existing soil and climatic conditions.

An informed decision can be made by the system using real-time weather data that has been provided by the OpenWeatherMap API. This will enable the platform to provide dynamic information of rainfall potential, changes in temperature and climatic conditions, which are also utilized to generate irrigation recommendations and assess any potential hazards of weather.

The system has many functional modules integrated in a single platform along with crop recommendation. They are soil health reports, yield predictions, market price, irrigation, sustainability and pests and disease prediction. These modules may be integrated to assist farmers to establish the profitability of crops, balance nutrients in the soil and arrange the agricultural activities in a more efficient way.

All the outputs and recommendations are presented as an interactive dashboard to visualize the crop suggestions, environmental status, yield forecasting and financial reports in the form of charts and graphical representations. It ensures ease of use and ensures the easy interpretation of the results by farmers who do not need technical expertise.

Overall, the suggested system provides a complete decision-support environment, with the combination of machine learning, live data, and easy-to-use visualization, bridging the divide between the innovative farming technology and the actual farming practice.

4. System Architecture

The Smart Crop Advisory System proposed has a modular and layered architecture in such a way that it will be scaled and enhanced data processing and integration of different components. The architecture has four key layers namely the user interface layer, application processing layer, machine learning layer and the data integration layer.

The user interface layer provides web based interfaces though the aid of which farmers can interface with the system. In order to be reactive and practical, it is written in the most recent web technologies, such as HTML, CSS, Bootstrap, and JavaScript. This layer gives an opportunity to input the parameters of the soil and view the recommendations and examine the results visually.

Processing layer is an application processing layer which was created using the Flask framework as the primary backend of the system. It receives user requests, incoming information, accomplishes the process of communication between modules, and provides smooth communication between the front and back-end components. This layer also equilibrates the process of the implementation of other system modules as crop recommendation, yield prediction, irrigation advisory and climate risk assessment.

The machine learning layer is the one that is tasked with predictive analytics. It is a trained random forest machine learning model that involve farm conditions

(nitrogen N, phosphorus P, potassium K, pH, temperature, humidity and rainfall). The model generates projections of the correct crops and expected yield or harvest based on trends on historical agricultural records. Ensemble learning has greater prediction accuracy and strength.

Data integration layer connects the system to external and external data sources. The inputs of users, history of predictions, and the system outputs are stored in a SQLite database. There are also external APIs like OpenWeatherMap that are used to retrieve current weather information like temperature, humidity, probability of rainfall and wind speed. This dynamic information increases the flexibility of the system in dynamic environmental conditions.

In general, the layered architecture allows communication between the components of the system to be performed efficiently and allows live data, machine learning predictions, and user interactions to be smoothly integrated into one decision-support platform.

System Architecture Diagram.

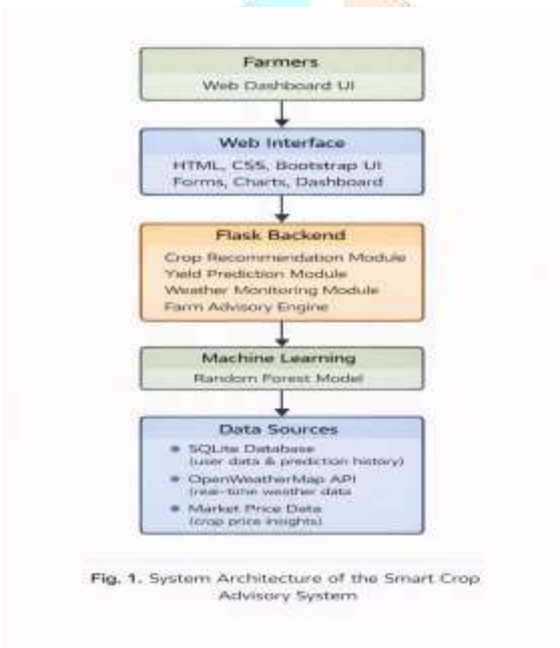

Fig. 1. Smart Crop Advisory System System Structure.

5. Methodology

The proposed Smart Crop Advisory System approach is a systematic pipeline process that incorporates data gathering, preprocessing, machine learning forecasting, live data incorporation, and advisory. It is expected to provide accurate farm management and crop recommendations based on soil and environmental conditions.

5.1 Data Input

The system receives input parameters via a web based interface. The required information provided by farmers is the required soil and environmental data including nitrogen (N), phosphorus (P), potassium (K), soil pH, temperature, humidity and rainfall.

These parameters are important indicators of soil fertility, climatic conditions, which directly affect crop growth and crop productivity. The dashboard is easy to use and allows farmers to enter data easily.

5.2 Prediction based on machine learning.

The system comprises a random forest classification model which was trained using agriculture based datasets of soil and environmental factors. Random Forest is chosen because it is more robust, it can operate under nonlinear relationships and also it can perform the best in agricultural predictions tasks [1], [12]. The trained machine learning model executes the input feature and calculates the right crops based on the trained patterns.

5.3 Weather Data Integration

Real-time weather data is integrated with the OpenWeatherMap API in the system, which provides a more accurate and flexible prediction. The temperature, humidity, rainfall, and the wind speed, are all environmental parameters, which are constantly retrieved and fed into the system.

The system facilitates such an integration that it offers climate-sensitive guidance, such as irrigation planning, climate risk assessment and environmental monitoring.

5.4 Advisory Generation

The system generates elaborate agricultural advice on the foundation of machine learning forecasts and real-time environmental measurements.

These include:

Crop recommendations

Predicted crop yield

Irrigation advisory

Soil health insights

Climate risk alerts

Profit estimation

Such outputs inform farmers about the conditions of agriculture and make informed decisions.

5.5 System Workflow

The system workflow may typically be split into input (data), preprocessing, prediction and advisory generation. The machine learning system processes soil and environmental information and produces information on crop and farm management.

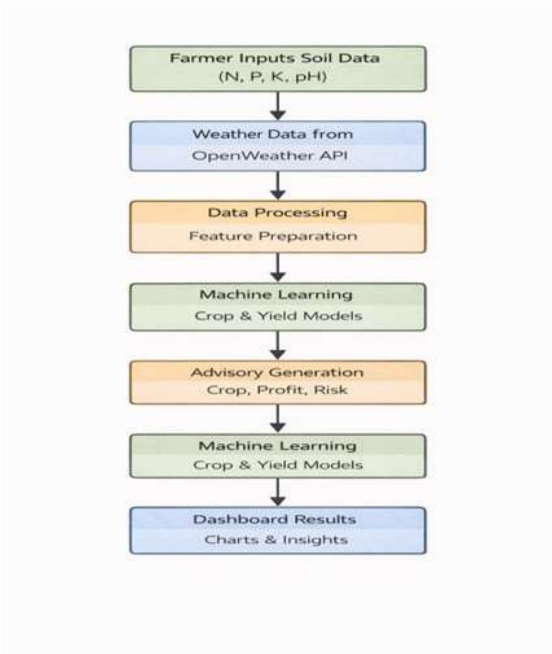

Fig. 2: Live Action of the Smart Crop Advisory System.

5.6 Dataset Description

The system makes use of a dataset on crop recommendations that contains the soil and environmental data of nitrogen (N), phosphorus (P), potassium (K), temperature, humidity, pH, and rainfall and its associated crop labels.

The data is saved in the CSV format and is manipulated with the help of Pandas library in Python. The preprocessing involves missing data and normalizing data of features in order to enhance the performance of the model.

The data are split into a training and a testing group of 80:20 and predictive performance of the model is evaluated.

Implementation

The Smart Crop Advisory System is applied as an agricultural decision support system web based platform with Python and Flask as the backend development, and HTML, CSS, Bootstrap, and JavaScript as the frontend development. The system has a user interface that is interactive and responsive to allow farmers to enter data and see the results.

The backend is in charge of the processing of user inputs, the logic of the application, and machine learning model integration. Two modules, crop recommendation and yield prediction, are applied with the help of Scikit-learn library, in which a random forest algorithm is applied to conduct a classification and a prediction task. The trained models are saved in the form of a serialized file and loaded to the flask application to make real-time predictions.

The system combines the open weather map real-time weather data through the OpenWeatherMap API. Environmental factors like temperature, humidity and rainfall are dynamically fetched and combined with soil inputs to enhance accuracy of prediction.

A SQLite database is used to store user inputs, prediction history and system outputs.This allows effective data control and acquisition of data to be analyzed and used in the future.

The platform has various modules such as crop recommendation, yield forecasting system, soil analysis, irrigation control, climate risk analysis, pests and disease identification, sustainability indicators, and market price evaluation. These modules are cooperative in the provision of specific agricultural information.

All the results are outlined in the form of charts and graphs depending on the interactive dashboard. This enhances usability and allows the farmers to make sound decisions by just interpreting the system outputs.

The system interface adopted is presented in Figure 3 according to which farmers can check their crop recommendation, yield forecast, and soil health information and farm data through a single dashboard.

Fig. 3. Smart Crop Advisor Dashboard Smart Interface.

7. Results and Discussion

The Smart Crop Advisory System was tested on different combinations of soil parameters and environmental parameters to determine its functionality in the real-world agricultural settings. The system was also effective to produce crop recommendations, yield forecasts and farm management insights via an interactive dashboard. On the test data, the model of the Random Forest had the predicted accuracy of about 94 percent, which proves to be a good predictive instrument in a task of crop recommendations [12]. These findings show that AI-based models could be important in enhancing agricultural decision-making because machine learning models can effectively determine the right crops to be planted under specific environmental factors [11].

The crop recommendation module reviews the major parameters in inputs, such as soil nutrients (N, P, K) at a pH level, temperature, humidity, and rainfall to identify the most appropriate crops to be grown. Random Forest is a suitable model that can be used to determine the best crops under such conditions. Given a sample input the system suggested Chickpea as the best crop to be grown after which the system recommended Grapes and Mango as alternatives.

Figure 4 shows the crop recommendation result with the confidence score, expected yield and return of investment (ROI). The outputs allow farmers to make wise decisions that are guided by productivity and profitability

.

Fig. 4. The results of the System in drawing its Crop Recommendation.

The system is also able to offer visual analytics to facilitate the decision-making process. The analytics module compares profitability of recommended crops and offers the indications of sustainability which include the soil health score, water efficiency, organic use, and biodiversity impact. These pieces of information assist farmers to consider the economic and environmental factors of crop choice. In figure 5, the profit comparison and sustainability metrics dashboard is presented.

Fig. 5. Comparison of Profitability and sustainability metrics Dashboard.

Besides that, the system includes real-time environmental monitoring, including the weather and soil visualization modules. Temperature, humidity, and rainfall are shown on the weather overview, whereas, on the soil nutrient visualization, the distribution of nitrogen, phosphorus, and potassium is presented. These images can help farmers to know the environmental conditions that influence the growth of crops. The dashboard of the weather and soil analysis is shown in Figure 6.

Fig. 6. Soil and Weather risk and opportunity chart.

The system also contains sophisticated advisory modules which include climate risk assessment, irrigation management, pest and disease detection and waste management analysis modules. The climate risk module analyses the possible risks, including drought, flood and heat stress, in accordance with weather conditions, and therefore farmers can take preventive measures. The climatic risk analysis interface is shown in figure 7.

Fig. 7. Climate Risk Assessment Module.

The system also offers a weather forecasting system that shows a 7-day forecast based on real-time data of the OpenWeatherMap API. The forecast entails the parameters of temperature trends, speed of wind, visibility and UV index. This attribute aids farmers to organize irrigation and agricultural practices. The weather forecast dashboard is depicted in figure 8.

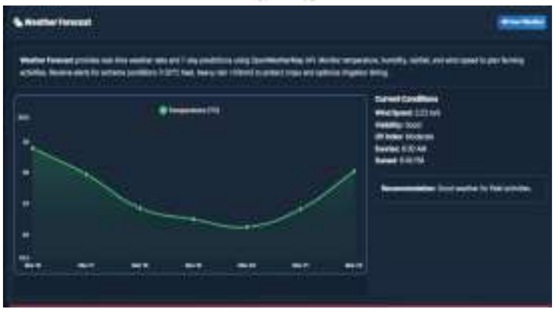

Fig. 8. Weather Forecast Dashboard

In addition, the system has other features that can be compared to crops, historical predictions, and sustainability assessment. The modules provide farmers with a comparison between the various crop alternatives depending on the yield potential, profitability, and environmental impact. The waste management module offers suggestions on the use of crop residues in compost production, biogas generation, and generation of organic fertilizers, as a way of ensuring sustainable agricultural production.

In general, the findings indicate that the suggested system can be successfully used to combine machine learning with live data analysis and visualization technologies to facilitate the use of data in agricultural decision-making. The system does not only increase the accuracy of crop selection but also increases the farm planning, risk management and sustainability.

Conclusion

The Smart Crop Advisory System is an efficient solution to the field of precision agriculture, offering the capacity to combine machine learning with real-time environmental information. It is a system that interprets soil and climatic variables to provide precise crop advisory and yield forecasts to enable farmers to make well-informed and evidence-based decisions. The Random Forest model was also very efficient as it was able to predict the use of individual crops in different conditions with high accuracy.

In addition to crop recommendation, the system possesses other options founded on weather-based advisory, soil analysis, irrigation planning, and sustainability information through an interactive dashboard. These characteristics allow the management of the farm, their resources more effectively as well as reduce risks in farming. Generally, the proposed system is a useful and scaled-up technology to apply in the modern farms, increased productivity and sustainable agricultural application.

References

[1] T. van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: A systematic literature review," *Computers and Electronics in Agriculture*, vol. 177, p. 105709, 2020.

[2] J. Cowls, A. Tsamados, M. Taddeo, and L. Floridi, "The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations," *AI & Society*, vol. 38, no. 1, pp. 283–307, 2021.

[3] P. Mittal, "A comprehensive survey of deep learning-based lightweight object detection models for edge devices," *Artificial Intelligence Review*, vol. 57, no. 9, 2024.

[4] D. B. Olawade, O. Z. Wada, A. O. Ige, B. I. Egbewole, A. Olojo, and B. I. Oladapo, "Artificial intelligence in environmental monitoring: Advancements, challenges, and future directions," *Hygiene and Environmental Health Advances*, vol. 12, p. 100114, 2024.

[5] T. L. Narayana, C. Venkatesh, A. Kiran, C. B. J, A. Kumar, S. B. Khan, A. Almusharraf, and M. T. Quasim, "Advances in real time smart monitoring of environmental parameters using IoT and sensors," *Heliyon*, vol. 10, no. 7, e28195, 2024.

[6] H. Hamdaoui, H. A. Kaddouri, I. Hamdi, F. Chkird, M. Derouiche, A. Alla, and N. Kouddane, "Toward Sustainable Farming: Artificial Intelligence Applications in the Detection and Management of Crop Diseases and Pests: A Systematic Review," *Basrah Journal of Agricultural Sciences*, vol. 38, no. 2, pp. 527–569, 2025.

[7] B. Jin and X. Xu, "Wholesale price forecasts of green grams using the neural network," *Asian Journal of Economics and Banking*, vol. 9, no. 3, pp. 463–490, 2024.

[8] S. Ramesh and D. Vydeki, "Recognition and classification of paddy leaf diseases using optimized deep neural network with Jaya algorithm," *IEEE Access*, vol. 7, pp. 132664–132677, 2019.

[9] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, and E.-H. M. Aggoune, "Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk," *IEEE Access*, vol. 7, pp. 129551–129583, 2019.

[10] J. Jeong, J. P. Resop, N. D. Mueller, D. H. Fleisher, K. Yun, E. E. Butler, and D. Kim, "Random forests for global and regional crop yield predictions," *PLoS ONE*, vol. 11, no. 6, 2016.

[11] M. T. H. Khan, A. Robson, and J. M. S. Sattar, "A review of applications of machine learning in precision agriculture," *Computers and Electronics in Agriculture*, vol. 178, p. 105760, 2021.

[12] A. Khanna and S. Kaur, "Evolution of Internet of Things (IoT) and its significant impact in the field of precision agriculture," in *Proc. IEEE Int. Conf. Computing for Sustainable Global*

Development (INDIACom), 2019, pp. 562–567.