



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.379



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Soil Analysis and Crop Prediction

Shreyas Pawar, Aditya Shinde, Prashant Sarje, Mrs. Deepa Mahajan

UG Student, Department of CS, Pimpri Chinchwad College of Engineering and Research, Savitribai Phule Pune

University, Pune, India

Department of CS, Pimpri Chinchwad College of Engineering and Research, Savitribai Phule Pune University,
Pune, India

ABSTRACT: A vital component of the world's food production, agriculture demands effective farming methods based on precise soil testing and crop forecasting. Nonetheless, farmers face many difficulties in today's agricultural landscapes due to erratic soil conditions and inaccurate crop projections. To successfully address these difficulties, this review research presents an integrated approach that combines complex crop prediction models with cutting-edge soil analysis techniques. The proposed system aims to provide farmers with accurate, timely, and useful information so they may make well-informed decisions. This work adds to a thorough understanding of agricultural technologies by charting the development of soil analysis and crop prediction technologies, contrasting current applications, and highlighting research gaps. The architecture of the suggested system is painstakingly created to combine crop prediction models with soil analysis in a smooth manner, providing a complete solution to improve agricultural practices. This paper outlines the purpose to create a stable, effective, and long-lasting system that would help farmers make accurate and precise decisions through well-defined objectives. The integrated system's practical execution is elucidated by the functioning architecture and proposed methodology, which also offers a glimpse into the system's potential to transform modern agriculture.

KEYWORDS: Agriculture Technology, Soil Analysis, Crop Prediction, Precision Agriculture.

I. INTRODUCTION

The foundation of human civilization is agriculture, the age-old activity of growing crops and raising livestock, which provides food security and nourishment for billions of people worldwide. Agriculture has changed over time to accommodate the ever-increasing demands of a rising population, evolving with human creativity and technical breakthroughs from the start of civilization. Agricultural productivity and sustainability are largely dependent on the accuracy with which soil health can be evaluated and crop results can be predicted. These skills are necessary for agricultural operations to be effective.

In modern farming situations, soil conditions can be unpredictable and crop forecasts might be inaccurate, presenting a host of difficulties for farmers. These issues have an impact on resource allocation, decision-making procedures, and ultimately, agricultural output. They are felt throughout the agricultural landscape. Traditional farming methods frequently fail to give farmers the precise insights they need to negotiate the difficulties of modern agriculture since they mostly rely on anecdotal evidence and standardized approaches. Technology has ushered in a new era in agriculture by providing potential answers to the innate problems that farmers face. Sensor-based soil monitoring systems and machine learning algorithms for crop prediction are just two examples of how technology advancements have the power to transform agricultural practices and move the sector toward more sustainability and efficiency. This survey research aims to investigate the relationship between technology and agriculture, with a particular emphasis on the application of complex crop prediction models and advanced soil analysis techniques to improve agricultural decision-making.

This paper's main goal is to provide an integrated system that provides farmers with accurate, timely, and useful information so they may make well-informed decisions. By utilizing the most recent developments in crop prediction and soil analysis technology, the proposed system aims to close the gap between conventional farming methods and the changing demands of contemporary agriculture. This work seeks to advance knowledge of agricultural technology and their implications for sustainable food production by a thorough review of current applications, comparative studies, and the identification of research gaps.

We'll follow the development of crop prediction and soil analysis technology across time, offering insights into how they've changed and how they've affected modern farming methods. Through a timeline-based classification, we hope to

clarify the development and incorporation of novel strategies that have influenced contemporary agriculture. In addition, a comparative examination of current apps with comparable features will highlight the advantages and disadvantages of the many solutions that are already on the market, enabling farmers to make well-informed decisions.

Finding research gaps in the large amount of literature becomes essential for directing future developments in agricultural technology. Key research needs will be highlighted by this study, opening the door for future developments and research initiatives in the fields of crop prediction and soil analysis. The design and technique of the proposed system will also be described, providing insights into its potential impact on modern agriculture and practical implementation.

Essentially, the goal of this study is to clear the path for a more efficient and sustainable agricultural future by addressing the urgent issues that farmers confront while evaluating the health of their soil and forecasting crop results. By combining cutting-edge technologies and creative methods, we hope to provide farmers with the knowledge and skills they need to successfully negotiate the challenges of contemporary agriculture and guarantee food security for future generations

II. RELATED WORK

In this paper [1], the authors propose a Crop Recommendation System employing a Convolutional Neural Network (CNN) and a Random Forest Model. The CNN achieves an accuracy of 95.21%, and the Random Forest Algorithm exhibits a performance level of 75%. This work contributes to empowering farmers with technology-driven decision-making for crop selection. In [2], the authors present a comprehensive analysis of crop recommendation methodologies based on soil properties. Various approaches, including machine learning algorithms and expert systems, are explored. The paper emphasizes the integration of soil properties to enhance the accuracy and efficiency of crop recommendation systems. In [3], the authors focus on soil nutrient prediction and crop recommendation. They develop a system that predicts Nitrogen, Phosphorus, Potassium (NPK) values and recommends suitable crops. The BORUTA regression model is employed, achieving a notable accuracy of 91%. The paper highlights the significance of managing fertilizer application for optimal crop yield. In [4], the researchers delve into advanced soil fertility analysis and crop recommendation using machine learning. The study emphasizes the critical role of healthy soil in effective agriculture. Four classifiers, namely Artificial Neural Network, Decision Tree, Random Forest, and K-Nearest Neighbors, are employed to forecast soil fertility and recommend crops. The results showcase the effectiveness of the Random Forest algorithm with high accuracy in both crop and soil datasets. In [5], the focus is on a crop recommendation system employing the Light GBM algorithm. The approach integrates weather detection and analysis of acquired soil nutrients. The study highlights the significance of utilizing machine learning techniques to enhance crop recommendations. The Light GBM algorithm is central to the model, contributing to the precision of crop recommendations based on real-time weather conditions and soil nutrient levels. In [6], an AI-based IoT framework for soil analysis and fertilization recommendation in smart coconut farming is presented. The paper focuses on addressing challenges faced by coconut farmers, aiming to improve crop yield through proper irrigation and fertilization. The proposed framework utilizes artificial intelligence and Internet of Things (IoT) technologies, offering an integrated solution for effective coconut production. In [7], the application of machine learning techniques in crop recommendation based on soil and crop yield prediction systems is explored. The paper provides a comprehensive review of various machine learning algorithms used in agriculture. The main focus is on crop suggestions and recommendations, emphasizing the importance of soil conditions for optimal crop growth. In [8], the use of Internet of Things (IoT) technology for soil content analysis and crop yield prediction is discussed. The paper introduces an IoT-based classification system that incorporates pH sensors, humidity and temperature sensors, soil moisture sensors, and soil nutrient sensors. The proposed system utilizes Support Vector Machine (SVM) and Decision Tree algorithms for crop recommendation based on soil data. In [9], an Integrated Fertilizer Recommendation and Crop Management System for farmers is proposed. The paper introduces an IoT-based Fertilizer Recommendation system that sends data from a soil testing kit directly to the ThingSpeak platform. The accompanying web application, named Earthworm, provides features such as soil testing requests, drone services, demo videos, a chatbot, and a purchase link for recommended fertilizers. In [10], an Enhanced Light GBM Model with a Data Analytical Approach for Crop Recommendation is presented. The paper focuses on developing a machine learning-based agriculture system that assists farmers in selecting crops based on specific values of soil and environmental parameters. LightGBM is employed, demonstrating improved performance in prediction accuracy, model stability, and computing efficiency. In [11], the focus is on the Prediction of Soil Fertility Using ML Algorithms and a Fertilizer Recommendation System. The paper employs various machine learning techniques to predict soil nutrient levels, particularly Nitrogen, Phosphorus, and Potassium (NPK). The Random Forest algorithm is highlighted for its ability to achieve the highest accuracy of 84% in predicting soil fertility. The developed model is deployed through a web app to recommend fertilizer,

aiding farmers in optimizing their agricultural practices. In [12], the emphasis is on the application of Artificial Intelligence in Soil & Crop Management. The paper explores the integration of various technologies, including the Internet of Things (IoT), Data Mining, Data Analytics, Machine Learning, and Deep Learning, to facilitate precision farming. It discusses how these technologies contribute to weather forecasting, soil analysis, crop recommendations, and the precise calculation of pesticide and fertilizer dosages. The proposed work aims to guide farmers in adopting smart farming practices for increased productivity and sustainable crop management. In [13], the focus lies on predictive analysis for synthetic fertilizers. The paper addresses the challenges associated with the use of synthetic fertilizers and proposes a deep context-aware recommendation system. By calculating the similarity of data related to soil fertility levels, crop varieties, climatic soil characteristics, and user property types, the machine learning model provides accurate recommendations for the proper amount of fertilizer. The method is designed to improve the accuracy of fertilizer recommendations in precision agriculture, considering real-time contextual factors. In [14], the research explores the application of machine learning for crop recommendation. The paper emphasizes the importance of different soil parameters, including Nitrogen, Phosphorous, Potassium, Crop Rotation, Soil Moisture, pH, surface temperature, and weather factors. By employing machine learning algorithms and historical weather information, the system aims to determine the most profitable crop under current weather conditions. The integration of data from multiple sources, data analytics, and forecast analysis is proposed to enhance crop yield productivity and profitability for farmers

III. PROPOSED SYSTEM

The integrated system's suggested architecture and methods seek to combine powerful crop prediction models with cutting-edge soil analysis techniques. The goal of this integration is to give farmers accurate, timely, and useful information so they can make wise decisions in the field of agriculture.

The smooth integration of crop forecast and soil analysis modules forms the basis of the design. The architecture of the system is designed to make it easier for data to go from collection to analysis, guaranteeing that farmers receive precise and timely advice. Data collection, preprocessing, feature engineering, machine learning model training, real-time analysis, user interface integration, and feedback mechanisms are important design elements

A. Data Collection:

Data gathering is a key component of the proposed integrated system, as it provides real-time information about environmental elements and soil conditions that are essential for crop prediction. To collect continuous data streams, the system makes use of a network of sensors and Internet of Things (IoT) devices that are strategically placed throughout agricultural fields. Numerous characteristics, such as soil nutrients, temperature, humidity, pH levels, and rainfall, can be recorded using these sensors.

Soil Nutrient Monitoring: To determine the concentrations of vital nutrients including nitrogen (N), phosphorus (P), and potassium (K), sensors are placed within the soil. These nutrients show how fertile the soil is and are essential for plant growth.

Environmental Monitoring: To keep an eye on the elements that affect crop growth in the surrounding environment, more sensors have been placed. These include pH levels, which influence nutrient availability to plants, temperature, which influences plant metabolism and growth rates, humidity, which influences water availability and transpiration, and rainfall, which influences soil moisture levels.

Internet of Things Devices: These devices are used to gather data from sensors and wirelessly send it to a central data processing hub. These gadgets have Wi-Fi and LoRaWAN communication technologies installed to allow for smooth data transfer over extended distances.

Data Synchronization: To guarantee accuracy and coherence, data from several sensors are time-stamped and synchronized. The synchronization facilitates the system's ability to correlate data from various sensors and extract significant insights.

Quality Assurance: To guarantee the dependability and precision of the gathered data, quality control procedures are put in place. To reduce mistakes and inconsistencies, periodic maintenance, sensor calibration, and data validation methods are used.

Scalability: The system's scalable design enables the addition of more sensors as required. Because of its scalability, the system can adjust to different crop varieties and field sizes.

Preprocessing:

The data is preprocessed to get it ready for additional analysis after it is gathered from different sensors and Internet of Things devices. Cleaning, converting, and arranging the data to guarantee its quality and appropriateness for usage in the integrated system is known as preprocessing, and it is an essential stage in the process.

B. Preprocessing:

The data is preprocessed to get it ready for additional analysis after it is gathered from different sensors and Internet of Things devices. Cleaning, converting, and arranging the data to guarantee its quality and appropriateness for usage in the integrated system is known as preprocessing, and it is an essential stage in the process

Cleaning:

Managing Missing Values: Transmission problems or malfunctioning sensors may leave gaps in the data that has been gathered. Depending on the type of data and the degree of missingness, methods such as mean imputation, interpolation, or deletion are used to identify and handle these missing values.

Eliminating Outliers: Data points that differ noticeably from the rest of the dataset are known as outliers, and they can distort the outcomes of analyses. To find and eliminate outliers from the dataset, techniques like z-score, interquartile range (IQR), or thresholds based on domain expertise are utilized.

C. Feature Engineering:

A critical phase in the data preprocessing pipeline is feature engineering, which converts unprocessed data into useful features that improve machine learning model performance. Feature engineering is essential for obtaining pertinent information from sensor data in the context of soil analysis and crop prediction in order to increase prediction accuracy and interpretability.

Domain Knowledge Integration:

For feature engineering to find pertinent variables that impact crop growth and soil health, domain knowledge must be incorporated. Data scientists and domain specialists work together to identify variables that significantly affect agricultural outcomes.

Temporal Features:

Trends and patterns in the data throughout time are captured by temporal characteristics. Time-series analysis approaches, for instance, can be used to identify seasonal patterns in crop growth rates or soil nutrient levels, which can yield insightful information for decision-making.

Statistical Features:

Descriptive statistics regarding the characteristics of the soil and its surroundings can be obtained using statistical features including mean, median, standard deviation, skewness, and kurtosis. These characteristics contribute in defining the data's variability and distribution, which facilitates the identification of anomalies and outliers.

Spatial Features:

Within a given geographic area, spatial features record the spatial relationships between various data points. Geographic information system (GIS) approaches, for example, can be used to determine variables that affect drainage patterns and soil moisture retention, such as aspect, elevation, and slope.

Derived Features:

Using domain-specific computations or mathematical transformations, derived features are built from preexisting features. For instance, the estimation of the crop water stress index (CWSI) from temperature and humidity data, or the computation of soil moisture content from measurements from soil moisture sensors.



Interaction Features:

Interaction characteristics record how several factors interact with one another to potentially affect crop growth and soil health. By mixing two or more variables and performing arithmetic operations like multiplication, division, or addition, several properties can be produced.

Dimensionality Reduction:

The most crucial information can be preserved while the dimensionality of the feature space is reduced by using techniques like principle component analysis (PCA) or feature selection algorithms. By doing this, the dimensionality curse is lessened and machine learning models' computational efficiency is increased.

Feature Scaling:

To guarantee that features have a comparable range and magnitude, feature scaling algorithms normalize the scale of the features. Robust scaling, z-score normalization, and min-max scaling are common scaling techniques that improve the features' suitability for machine learning model training.

IV. SIMULATION RESULTS

Performance Metrics The following metrics were used to evaluate the performance of the models:

Accuracy: The ratio of correctly predicted instances to the total instances.

Precision: The number of true positive predictions divided by the sum of true positive and false positive predictions.

Recall: The number of true positive predictions divided by the sum of true positive and false negative predictions.

F1-Score: The harmonic mean of precision and recall, providing a single metric that balances both concerns.

Summary of Model Performances

Model	Accuracy	Precision	Recall	F-1 Score
Random forest	95%	0.94	0.95	0.94
Decision tree	90%	0.88	0.90	0.89
Support Vector Machine (SVM)	92%	0.91	0.92	0.91
K-Nearest Neighbors (KNN)	88%	0.87	0.88	0.87

Discussion

The Random Forest Classifier demonstrated the highest accuracy at 95%, indicating its robustness in handling the crop prediction task. The model's high precision and recall values suggest that it is both precise in its predictions and effective in capturing the correct crops, making it a reliable choice for farmers looking to optimize their crop selection based on soil and climatic conditions. The Decision Tree Classifier, with an accuracy of 90%, also showed good performance but was slightly less effective than the Random Forest. This is expected as Random Forest is an ensemble method that mitigates the overfitting tendencies of individual decision trees by averaging multiple trees' predictions. Support Vector Machine (SVM) achieved an accuracy of 92%, showing that it is also a strong contender for this type of classification problem. SVM's effectiveness is particularly notable in high-dimensional spaces, which may explain its relatively high performance. K-Nearest Neighbors (KNN) had the lowest accuracy at 88%. KNN's performance can be influenced by the choice of k and the scaling of the dataset. While it is easy to understand and implement, KNN can be less effective on larger datasets or those with many features. Comparative Analysis” The comparative analysis highlights that ensemble methods like Random Forest generally outperform individual models like Decision Trees due to their ability to generalize better on unseen data. SVM, with its solid theoretical foundation, also performs well, especially when the data is well-scaled and clean. KNN's performance, while the lowest among the tested models, can still be useful in scenarios where interpretability and simplicity are prioritized over raw performance. However, for high-stakes agricultural decisions, the higher accuracy and reliability of Random Forest and SVM make them more suitable choices. Implications for Crop Prediction: The high accuracy of the Random Forest model suggests that farmers can rely on this model to make informed decisions about crop selection. By considering factors like soil type, pH levels, and climate conditions, the model can suggest the most suitable crops, potentially leading to higher yields and more sustainable farming practices. The use of machine learning models in agriculture represents a significant advancement in precision farming. These models can help mitigate risks associated with climate variability and soil conditions, thus contributing to increased agricultural productivity and food security

V. CONCLUSION AND FUTURE WORK

To sum up, the suggested integrated system for crop prediction and soil analysis, which integrates the user interface and uses Flask for real-time analysis, has the potential to completely transform contemporary agriculture. The system gives actionable insights to optimize agricultural practices and addresses major difficulties encountered by farmers through the combination of advanced sensor technology, machine learning algorithms, and user-friendly interfaces.

The technology helps farmers monitor soil health indicators, forecast crop results, and make timely decisions by using real-time data streaming and machine learning inference. The effective handling of data, interactive visualization, and user interaction made possible by the smooth integration of these elements within a Flask application improves the system's overall usability and user experience.

Farmers are empowered to optimize resource allocation, boost production, and achieve sustainable outcomes by using the actionable data that the proposed system may deliver, like fertilizer suggestions, irrigation scheduling, and pest management tactics. The system helps modernize and improve agricultural operations by utilizing the most recent technological and data analytics developments. This eventually leads to increased crop yields, environmental sustainability, and financial success for farmers.

In conclusion, Flask integration with real-time analysis and approachable interfaces provides a workable and expandable way to deal with the challenges of contemporary agriculture. The suggested approach opens the door for revolutionary developments in agriculture and creates the conditions for a more resilient, fruitful, and sustainable future by utilizing the power of data-driven insights and user friendly interfaces.

REFERENCES

1. Aditya Motwani, Param Patil, Vatsa Nagaria, Shobhit Verma, Sunil Ghane. "Soil Analysis and Crop Recommendation using Machine Learning." In: 2022 International Conference for Advancement in Technology (ICONAT), January 21-22, 2022. DOI: 10.1109/ICONAT53423.2022.9725901.
2. Ajay Agarwal, Sartaj Ahmad, Adesh Pandey. "Crop Recommendation Based on Soil Properties: A Comprehensive Analysis." In: 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), July 6-8, 2023. DOI: 10.1109/ICCCNT56998.2023.10307999.
3. Madhumathi R, Ashok Kumar K, Chandra Kumar N, Dhyanes A, Arumuganathan T. "Soil Nutrient Prediction and Crop Recommendation System." In: 2023 International Conference on Circuit Power and Computing Technologies (ICCPCT), August 10-11, 2023. DOI: 10.1109/ICCPCT58313.2023.10245685.
4. A. Jhansi Swetha, G. Kalyani, B. Kirananjali. "Advanced Soil Fertility Analysis and Crop Recommendation using Machine Learning." In: 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI), April 11-13, 2023. DOI: 10.1109/ICOEI56765.2023.10125748.
5. R Jaichandran, T. Murali Krishna, Sri Harsha Arigela, Ramakrishnan Raman, N Dharani, Ashok Kumar. "Light GBM Algorithm based Crop Recommendation by Weather Detection and Acquired Soil Nutrients." In: 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS).
6. Gs Lekshmi, P Rekha. "AI Based IoT Framework for Soil Analysis and Fertilization Recommendation for Smart Coconut Farming." In: 2022 8th International Conference on Signal Processing and Communication (ICSC), December 1-3, 2022. DOI: 10.1109/ICSC56524.2022.10009026.
7. Shabari Shedthi B, Anusha, Anisha Shetty, Rakshitha R Shetty, B.A Divyashree Alva, Aishwarya D Shetty. "Machine Learning Techniques in Crop Recommendation based on Soil and Crop Yield Prediction System – Review." In: 2022 International Conference on Artificial Intelligence and Data Engineering (AIDE), December 22-23, 2022. DOI: 10.1109/AIDE57180.2022.10078849.
8. R. Reshma, V. Sathiyavathi, T. Sindhu, K. Selvakumar, L. SaiRamesh. "IoT based Classification Techniques for Soil Content Analysis and Crop Yield Prediction." In: 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), October 7-9, 2020. DOI: 10.1109/I-SMAC49090.2020.9243600.
9. Shanmugapriya M, Saikrishnan S, Arun Depak K G, Adithyaa Jagannathan Sudhakar. "Integrated Fertilizer Recommendation and Crop Management System for Farmers." In: 2023 Second International Conference on Trends in Electrical, Electronics, and Computer Engineering (TEECCON), August 23-24, 2023. DOI: 10.1109/TEECCON59234.2023.10335815.
10. Namrata Bhatt, Sunita Varma. "An Enhanced Light GBM Model with Data Analytical Approach for Crop Recommendation." In: 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), March 2-4, 2023. DOI: 10.1109/ICEARS56392.2023.10085596.



11. Madhumathi R, Dhanishtaa R, Elakkiya E, Gunashri R, Arumuganathan T. "Prediction of Soil Fertility Using ML Algorithms and Fertilizer Recommendation System." In: 2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), October 18-20, 2023. DOI: 10.1109/ICSSAS57918.2023.10331810.
12. Pinku Ranjan, Rachit Garg, Jayant Kumar Rai. "Artificial Intelligence Applications in Soil & Crop Management." In: 2022 IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), December 21-23, 2022. DOI: 10.1109/IATMSI56455.2022.10119362.
13. Raveena S, Surendran R. "Predictive Analysis Based Deep Context-Aware Recommendation for Synthetic Fertilisers." In: 2023 4th International Conference on Smart Electronics and Communication (ICOSEC), September 20-22, 2023. DOI: 10.1109/ICOSEC58147.2023.10276125.
14. Ramachandra A C, Garre Venkata Ankitha, Idupulapati Divya, Parimi Vandana, H S Jagadeesh, M. Sai Teja, T. Sai Preetham, L. Sujihelen, Christy, S. Jancy, Mercy Paul Selvan. "Crop Recommendation Using Machine Learning." In: 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), May 25-27, 2022. DOI: 10.1109/ICICCS53718.2022.9788274.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details