



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 5, May 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Prediction of Crops Using IOT and Machine Learning

Dr. Manjunatha Kumar B H¹, Manaswi M², Manjusri N³, Meghana R⁴, Mytreya H B⁵

Head of the Department, Department of Computer Science and Engineering, S J C Institute of Technology,
Chickaballapur, Karnataka, India¹

U.G. Student, Department of Computer Science and Engineering, S J C Institute of Technology, Chickaballapur,
Karnataka, India^{2,3,4,5}

ABSTRACT: Crop prediction plays a vital role in ensuring agricultural sustainability and food security. Traditional methods of crop prediction often rely on historical data and manual observations, which may not sufficiently capture the dynamic nature of agricultural environments. In recent years, the integration of Machine Learning (ML) techniques and Internet of Things (IoT) technologies has provided new avenues for enhancing the accuracy and efficiency of crop prediction models. This paper presents a novel approach to crop prediction by leveraging ML algorithms and IoT sensors to analyze key environmental factors including soil moisture, pH levels, temperature, and humidity. By integrating real-time data from IoT sensors installed in agricultural fields, our proposed model aims to provide more accurate and timely predictions of crop yield and growth patterns. In conclusion, the combination of ML and IoT technologies offers a promising framework for enhancing crop prediction accuracy and optimizing agricultural practices. By leveraging real-time data and advanced analytical techniques, our approach contributes to the sustainable development of agriculture and the enhancement of global food security.

KEYWORDS: Precision agriculture, Internet of things, Machine learning, Crop selection, Sustainability

I. INTRODUCTION

Agriculture plays a pivotal role in India's economy, serving as a primary source of employment and livelihood for a significant portion of its population. However, Indian farmers often grapple with a critical challenge: selecting the most suitable crops for their specific soil conditions. The consequences of improper crop choices can lead to reduced productivity and economic setbacks. In response to this issue, precision agriculture has emerged as a promising solution. Precision agriculture, as the name suggests, involves the precise and judicious application of resources such as water, fertilizers, and other inputs to crops at the right time and in the right quantities. This approach aims to enhance agricultural productivity, optimize resource usage, and ultimately increase crop yields. Given the diverse and varied agricultural landscape in India, it is imperative to provide farmers with accurate and tailored recommendations for crop selection to ensure the success of precision agriculture initiatives. To address this need for precise crop recommendations, this paper presents a Crop Recommendation System that leverages the power of machine learning techniques. The system relies on a wealth of data, including information collected from soil testing laboratories and expert guidance on crop selection. By harnessing the capabilities of machine learning, particularly through the utilization of ensemble models, this system seeks to offer highly accurate and efficient recommendations to Indian farmers. Crop prediction in agriculture is a compound process and numerous representations have been proposed and verified to this end. The problem calls for the use of assorted datasets, given that crop cultivation depends on various factors. These factors include temperature, humidity, soil pH, soil moisture, and so on. Among the various machine learning techniques that are being used in this field, this paper uses Naïve Bayes, Random forests, and Decision trees to build efficient and accurate.

II. LITERATURE SURVEY

In this section, previously performed studies on crop prediction using machine learning and IoT are done. The paper [1] is a study on developing a crop prediction model based on features of the agricultural situation using ML algorithms like KNN, Decision trees, SVM, Random Forest, and bagging. The model has achieved an accuracy of 73.22% with the Decision tree algorithm and 83.24% with the KNN algorithm.

The paper [2] focuses on a Crop Recommendation System to maximize crop yield. This uses Precision agriculture, a

recommendation system, an ensemble model, SVM, ANN, random tree, and NB classifier to give crop recommendations to farmers. This work would help farmers to increase productivity in agriculture, and prevent soil degradation in cultivated land.

The paper [3] is based on the Intelligent Crop Recommendation System. Decision Tree, K Nearest Neighbor (K-NN), Random Forest, and Neural Network are four machine learning algorithms that are used for classification in the proposed model. Accuracy with the Decision Tree Algorithm is 90.20% and with K-NN is 89.78%.

The referenced paper [4] doesn't investigate using both ANN and SVR, but rather focuses on ANNs. It explores the potential of using machine learning (ML) specifically artificial neural networks (ANNs) to improve the retrieval of soil moisture content (SMC) from satellite data. The study combines C-band and X-band data, along with simulated compact polarimetry (CP) parameters to train the ANN. Their findings demonstrate that ANNs are a promising tool for estimating SMC, especially when used in conjunction with CP parameters.

In paper [5] is a study of arable farming systems in temperate regions, the paper aims to provide a broad understanding of diverse farming methods used in temperate climates.

III. PROPOSED SYSTEM

The proposed system consists of three main components: data acquisition through IoT sensors, data processing and analysis using ML algorithms, and decision support for farmers. IoT sensors are strategically deployed across agricultural fields to collect real-time data on soil moisture, pH levels, temperature, and humidity. This data is transmitted to a centralized system for processing and analysis. In the data processing phase, ML algorithms are employed to analyze the collected data and develop predictive models for crop yield estimation. Various ML techniques, including ensemble, Random Forests, and Decision trees, are explored to capture the complex relationships between environmental factors and crop performance. Once the predictive models are trained, the system provides decision support to farmers by generating timely recommendations for irrigation scheduling, fertilizer application, and other crop management practices. These recommendations are tailored to the specific needs of each agricultural field based on real-time environmental data.

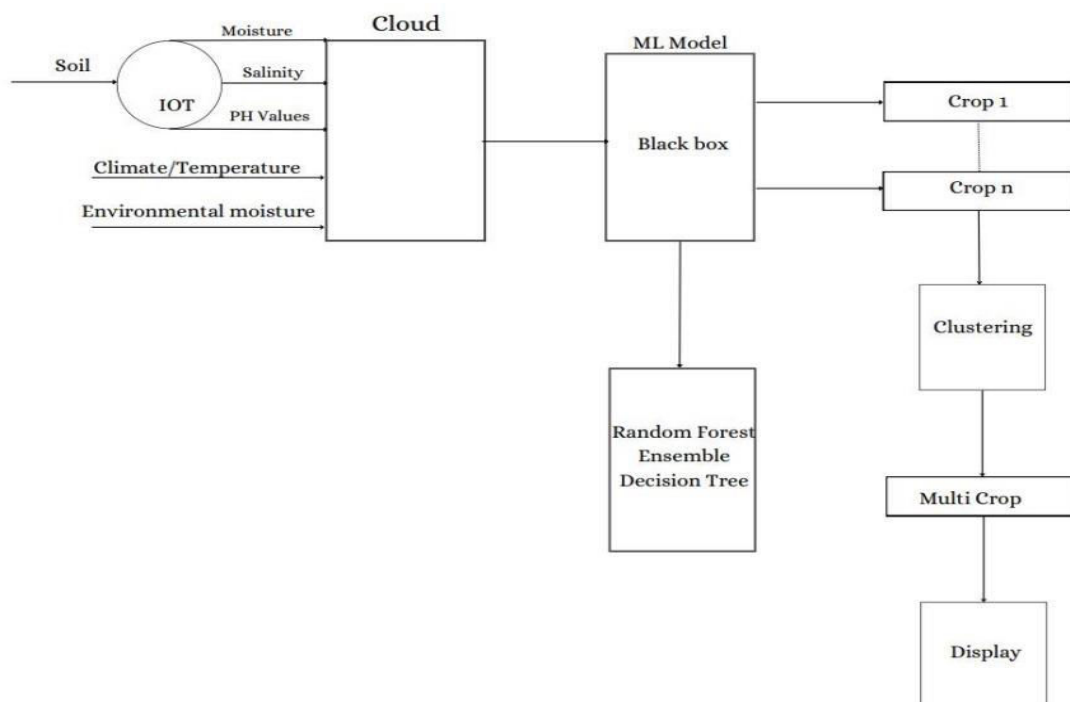


Fig 1: Architecture of the proposed system

IV. METHODOLOGY

Collection of Data: The success of any machine learning algorithm hinges on two crucial factors: the number of parameters it can handle and the quality of the training data. The "India Agriculture and Climate Data Set" serves as the foundation, offering a wealth of historical information spanning 30 years (1957-1986). This comprehensive dataset covers 13 major Indian states and over 270 districts, providing insights into a vast geographical area.

But the richness goes beyond location. The data encompasses information on 20 crops, including 5 major ones (bajra, jowar, maize, rice, and wheat) and 15 minor crops. This allows the model to learn nuances of various agricultural practices. It delves into critical factors influencing crop growth.

Soil Moisture: Precise data on soil moisture empowers farmers with targeted irrigation practices. This prevents water waste and ensures crops receive the right amount of water at different stages, promoting efficient water management and reducing production costs.

Soil pH: Knowing the soil pH allows farmers to take corrective measures if it falls outside the ideal range for a desired crop. By adding amendments like lime or sulphur, they can create a more favourable environment for optimal growth.

Climate Data: The dataset includes monthly humidity and temperature data, providing insights into the overall climate patterns that crops experience. By incorporating this rich tapestry of historical data on soil, weather, and diverse crops, the machine learning model can learn complex relationships between these factors. This knowledge can then be harnessed for tasks like predicting crop yields or recommending optimal farming practices, ultimately empowering farmers to make data-driven decisions and improve agricultural outcomes.

Data Preprocessing To prepare data for machine learning, this process tackles two key issues. First, missing values (represented by ".") are replaced with large negative values. This helps the model identify them as outliers. Second, since supervised learning requires labeled data, new class labels (0 or 1) are created based on crop yield (production divided by area). Crops with a higher production-to-area ratio are labeled 1 (potentially higher yield). This data preprocessing ensures the machine learning model can work with clean and informative data for better results.

V. ALGORITHMS

Naive Bayes: This algorithm predicts outcomes based on probabilities. It assumes features (data points) are independent and calculates the likelihood of a specific outcome based on evidence. It's fast, easy to implement, and works well with various data types. However, it struggles when features are interrelated, which is common in realworld data. Naive Bayes, a machine learning workhorse, utilizes the power of probability for classification tasks in agriculture. Imagine a farmer pondering the best crop for their land. Naive Bayes steps in, analyzing various factors like weather patterns, past crop performance, and market prices. Based on this data, it calculates the likelihood of success for different crops, guiding the farmer towards a data-driven decision. It is shown in Equation 1.

$$P(y | X) = P(X | y) \cdot P(y) / P(X) \quad (1)$$

where $P(y | X)$ = Posterior probability, $P(X | y)$ = Likelihood, $P(X)$ = Evidence $P(y)$ = Prior probability.

By analysing historical data and current trends, Naive Bayes empowers farmers to make informed decisions about crop selection. This data-driven approach can significantly improve agricultural outcomes, leading to better yields and increased income, especially in a world facing climate change. Overall, Naive Bayes offers a user-friendly and powerful tool for agricultural classification. Its versatility and probabilistic foundation make it a valuable ally for farmers seeking to navigate the complexities of crop selection in a data-rich world.

Decision Trees: these are a cornerstone of supervised machine learning, offering a powerful and interpretable approach to classification and prediction tasks. Imagine a flowchart – that's essentially what a DT is, but specifically designed for machine learning. It uses a tree-like structure where each branch represents a decision point, and the path you take through the tree leads you to a final prediction. DTs excel at their simplicity. Every internal node in the tree represents a question (test) asked about the data's attributes (features). Each branch represents a possible answer, and the leaf nodes at the end of the branches hold the final class (category) that the data point belongs to. This intuitive structure makes DTs easy to understand and visualize, even for people without a deep machine-learning background.

There are two main types of DTs: categorical and continuous. Categorical DTs are used when the target variable you're trying to predict has distinct categories (e.g., classifying an email as spam or not spam). Continuous DTs, on the other hand, come into play when the target variable has numerical values (e.g., predicting the price of a house).

However, DTs also have limitations. Since they rely on a series of sequential decisions, they can be susceptible to noisy data or irrelevant features. Additionally, building very complex DTs can lead to overfitting, where the model performs well on the training data but poorly on unseen data. Despite these limitations, DTs remain a popular choice for various reasons. In conclusion, Decision Trees offer a powerful and easy-to-understand approach to classification and prediction in machine learning. By understanding their structure, types, and attribute selection methods, you can leverage this fundamental tool for various applications.

Random Forest: Random Forest (RF) stands tall as one of the most successful supervised learning algorithms. Its secret lies in the power of collaboration. This approach, known as ensemble learning, tackles complex problems by combining the strengths of multiple classifiers, ultimately enhancing the model's performance. Imagine a vast forest – that's essentially what RF builds. Each tree within the forest represents a decision tree, a popular machine-learning model that uses a series of yes-or-no questions to classify data or predict outcomes. What sets RF apart is the way these trees are grown. Here's the twist: randomness is key. When constructing each tree, RF randomly selects a subset of features (data points) to consider at each decision split. This injects a healthy dose of diversity into the forest. Trees become less correlated with each other, preventing them from all getting stuck in the same rut and potentially overfitting the data. Overfitting occurs when a model becomes too specialized to the training data and performs poorly on unseen data. So how does the forest make a decision? Each tree votes! In classification tasks (e.g., spam email or not?), the majority vote wins. For numerical predictions (e.g., house price), the average of all three predictions is taken. This democratic approach leads to more robust and accurate predictions compared to a single decision tree. The benefits of Random Forest are Simplicity Versatility, Big Data Friendly, and Overfitting Fighter. However, no algorithm is perfect. Here are some limitations to consider Resource Hungry, Sparse Data Woes, Limited Extrapolation. Despite these limitations, Random Forest's accuracy, ease of use, and ability to handle various tasks make it a popular choice for a wide range of applications. From predicting pest attacks in agriculture to optimizing irrigation systems, Random Forest empowers us to leverage the power of collective intelligence in the machine-learning world.

VI. RESULTS

All implementations are carried out on Windows 10 having hardware configuration of Intel core i5 processor with 8GB internal RAM and 1.60GHz of CPU speed. The hardware setup of the project is given below, where the components used in the proposed work are mentioned as, Arduino UNO with Atmega processor, Dht11 temperature and moisture sensor, the soil moisture, and the 16*2 display.

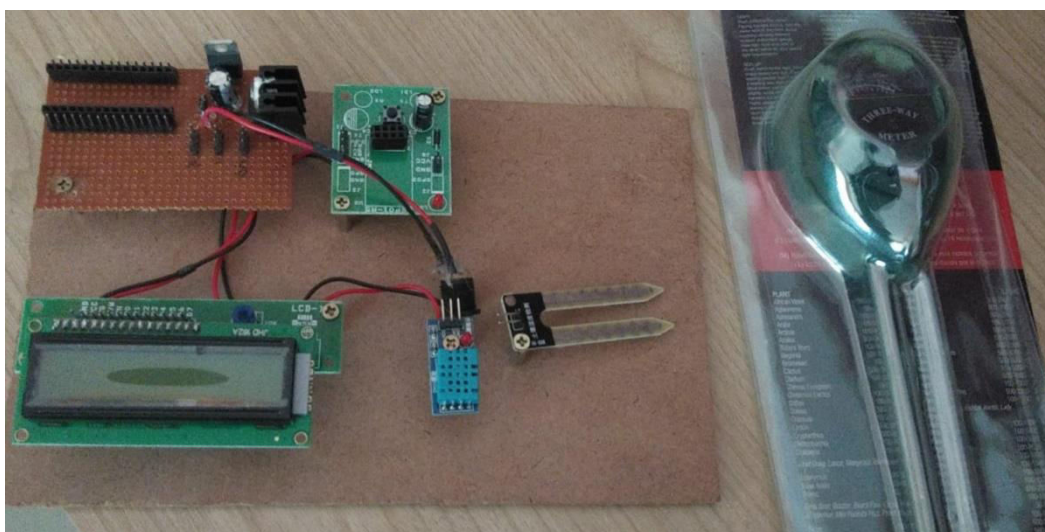


Fig 2: IoT setup of the crop prediction model

IoT sensors are deployed to collect real-time data on soil moisture, pH levels, temperature, and humidity. Collected data from the IoT sensors is transmitted securely to cloud storage platforms for centralized processing and analysis. Cloud storage solutions are utilized to efficiently store and manage the large volume of sensor data collected over time. Cloud-based storage and computing solutions offer cost-effective alternatives to traditional on-premises infrastructure, minimizing hardware investment and operational expenses. The efficiency of cloud-based data storage and processing will accelerate data analysis and decision-making, enhancing overall agricultural productivity.

A user-friendly interface for the website is developed to visualize real-time sensor data and provide actionable insights to farmers. The UI is seamlessly integrated with decision support systems, enabling farmers to receive real-time recommendations based on sensor data analysis.



Fig 3: The front-end of the crop prediction

The results of the IoT implementation and UI design will demonstrate the successful deployment of sensors and the development of a user-friendly interface for real-time data analysis. The integrated system offers valuable insights into soil conditions and environmental factors, empowering users to optimize crop management practices and enhance agricultural productivity.

VII. CONCLUSION

This paper proposes a beneficial system for Indian farmers that combines a Crop Suitability Predictor and a Rainfall Predictor. The Crop Suitability Predictor leverages environmental and geographical data to recommend suitable crops, while the Rainfall Predictor offers valuable insights for water management and planting schedules. Both models boast high accuracy, making them dependable tools for informed decision-making. While the current system excels at recommending individual crops, future iterations can be even more powerful by incorporating features for optimizing crop combinations. Also provides the recommendation for multi-crop. Overall, this research has the potential to revolutionize agricultural practices in India. By providing farmers with a comprehensive decision-making tool that considers both environmental and economic factors, this system can lead to increased yields, improved resource management, greater profitability, and a more sustainable agricultural future for India.

REFERENCES

- [1] S. P. Raja, Barbara Sawicka, Zoran Stamenkovic, (Senior Member, IEEE), and G. Mariammal Crop Prediction Based on Characteristics of the Agricultural Environment Using Various Feature Selection Techniques and Classifiers, 2022.
- [2] Rohit Kumar Rajak, Ankit Pawar, Mitalee Pendke, Pooja Shinde, Suresh Rathod, Avinash Devare Crop Recommendation System to Maximize Crop Yield using Machine Learning Technique, 2017.
- [3] Zeel Doshi, AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms,

2018.

- [4] E.Santi, M.Daboor, S.Pettinato, S.Paloscia, C.Notarnicola, F. Greifeneder, G. Cuzzo, on the use of machine learning and polarimetry for estimating soil moisture from radarsat imagery over italian and canadian test sites, 2019.
- [5] Jürgen Ehrmann & Karl Ritz, Plant: soil interactions in temperate multi-cropping production systems, 2014.
- [6] Mina Nath Paudel, Multiple Cropping for Raising Productivity and Farm Income of Small Farmers, 2016.
- [7] Fariha Shahrin, Agricultural Analysis and Crop Yield Prediction of Habiganj using Multispectral Bands of Satellite Imagery with Machine Learning, 2020.
- [8] Katharina Waha, Multiple cropping systems of the world and the potential for increasing cropping intensity, 2020.
- [9] G. Shivnath and S. Koley, “Machine learning for soil fertility and plant nutrient management using back propagation neural networks,” Int. J. Recent Innov. Trends Comput. Commun., vol. 2, no. 2, pp. 292–297, 2014
- [10] S. Pudumalar, E. Ramanujam, R. R. Harine, C. Kavya, T. Kiruthika, and J. Nisha, “Crop recommendation system for precision agriculture,” in Proc. 8th Int. Conf. Adv. Comput. (ICoAC), 2017, pp. 32–36.
- [11] “INDIA AGRICULTURE AND CLIMATE DATA SET”, available at https://ipl.econ.duke.edu/dthomas/dev_data/datafiles/india_agric_climate.htm, visited in November 2017.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details