International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

Crop prediction and fertilizer recommendation system using machine learning and IoT

Atharva Raut[1], Bhavy Mittal[2], Mayur Patel[3] , Vedant Polshettiwar[4], Prof. Sagar Padiya[5]

[12345]*Department of Information Technology, SSGMCE, Shegaon, Maharashtra*

Abstract - Agriculture is of utmost importance for survival, and machine learning (ML) and IoT can provide valuable solutions for crop selection issues. Therefore, this study aims to predict crops suitable for farmers' soil by utilizing various ML and IoT techniques. The classification model employed is Random Forest as its showing the highest accuracy. The application of ML algorithms to predict crops and suggest which fertilizers to use can aid farmers in making decisions regarding which crops to grow based on factors such as NPK values, humidity, temperature, and rainfall thereby bridging technology and the agricultural sector.

***Key Words*:** soil moisture sensor, npk sensor, random forest

1.INTRODUCTION

Agriculture occupies 60% of India's landmass and provides for the fundamental necessities of 1.4 billion individuals. For the well-being of the farmers, modernization of agriculture is carried out today. The production of crops is heavily reliant on weather conditions, environmental shifts, rainfall (which can often be unpredictable), water management, and the use of pesticides. This makes it difficult for farmers to achieve their desired crop. To address this issue, researchers have turned to machine learning techniques to improve crop yield and quality.[2]

Machine learning is a method of teaching machines without the need for explicit programming, it improves machine execution by portraying the consistency and pattern of information. In this work various machine learning methods such as Decision Tree, Naïve bayes, Logistic Regression, and Random Forest were employed to predict crops for different states.[7]

Since agriculture is the most dominant activity across all cultures and civilizations throughout history. It's not only a significant aspect of the growing economy but is also crucial for the survival and the future of humankind. It is a major contributor to employment and has seen an exponential increase in production requirements over time. However, people are misusing They can also be used in landscaping and gardening to conserve water and avoid over-watering. Moreover, soil moisture sensors are extensively used in research to monitor variations in soil moisture over time and analyze the correlation between soil moisture and plant growth. Soil moisture sensors operate within a voltage range of 3.3 to 5 volts.
technology to produce mass quantities, resulting in the creation of new hybrid varieties that do not provide the essential contents found in naturally produced crops, ultimately leading to environmental damage. Most of these unnatural techniques are employed to avoid losses, but with accurate information on crop, producers can minimize losses.[7] Machine learning, a rapidly growing approach, is spreading to every sector and helping make viable decisions to maximize its applications. Most gadgets today are worked with by models examined before the arrangement. The primary objective is to use machine learning models to boost agricultural sector throughput. Due to the relatively high number of parameters, the prediction is influenced by the knowledge provided during the training period. Precision agriculture, which prioritizes quality over unfavorable environmental factors, is the primary focus. Different machine learning classifiers, like logistic regression and random forest, are used to find a pattern in order to make accurate predictions and combat the erratic trends in temperature and rainfall. We came to the conclusion that the random forest algorithm provides the most accurate value by utilizing the aforementioned machine learning classifiers.[8] By collecting historical data on temperature, weather, and other factors, the system predicts which crops will be grown. The application runs the calculation and showcases the rundown of reasonable harvests for the entered information with the anticipated yield.

2. Methodology

The system relies on sensors to sense the soil parameters in real time and use an external dataset. The data obtained in real-time is stored on a local machine, and machine learning algorithms are employed for subsequent analysis.

A. Soil Moisture Sensor

A device designed to measure the moisture content of soil is known as a Soil Moisture Sensor. It is an electronic device that calculates the amount of moisture present in the soil.

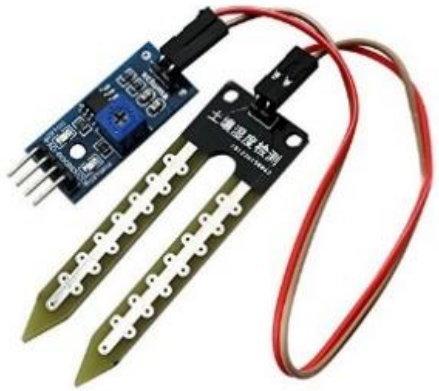

Figure 1: Soil Moisture Sensor

It typically comprises a probe that is placed in the soil to gauge its moisture content based on its electrical

conductivity. Soil moisture sensors are widely employed in agriculture to optimize irrigation practices and ensure appropriate water supply for crops.
They can also be used in landscaping and gardening to conserve water and avoid over-watering. Moreover, soil moisture sensors are extensively used in research to monitor variations in soil moisture over time and analyze the correlation between soil moisture and plant growth. Soil moisture sensors operate within a voltage range of 3.3 to 5 volts.

B. Soil NPK Sensor

A Soil NPK Sensor is a device that detects the levels of nitrogen, phosphorus, and potassium (NPK) in the soil. These are the primary nutrients necessary for plants' healthy growth and development. Soil NPK sensors use different methods, such as electrochemical or spectroscopic techniques, to measure NPK concentration in the soil. The usage of soil NPK sensors is crucial in agriculture because it helps farmers optimize fertilization practices, reduce fertilizer waste, and enhance crop yields. By determining the NPK levels in the soil, farmers can apply the right amount of fertilizer at the right time, preventing under or over-fertilization that can lead to plant stress or environmental pollution. Besides agriculture, soil NPK sensors can also be employed in research and environmental monitoring to comprehend nutrient cycles in soil and assess soil health. With this information, land managers and researchers can develop strategies to maintain soil fertility, prevent nutrient depletion, and restore degraded soils. Overall, soil NPK sensors are a critical tool for sustainable agriculture and environmental management, providing valuable data for precision farming and resource conservation. Soil NPK (Nitrogen, Phosphorus, and Potassium) sensors measure the levels of these essential nutrients present in the soil. The specific values and range of soil NPK sensor will depend on the its operating requirements.

Figure 3: Soil NPK Sensor

Typically, the measurement range of soil NPK sensors varies based on the nutrient being measured. The following are the general measurement ranges for soil NPK sensors:

- Nitrogen (N): 0-2000 mg/kg
- Phosphorus (P): 0-1000 mg/kg
- Potassium (K): 0-2000 mg/kg

The values measured by a soil NPK sensor may also be affected by soil type, pH, and other environmental factors. Calibration of the sensor based on local conditions may be required to obtain accurate measurements.
It's important to note that soil NPK sensors provide an estimate of nutrient levels in the soil, and laboratory analysis may be necessary to confirm the measurements. Additionally, the interpretation of the sensor readings requires knowledge of the crop being grown and its specific nutrient requirements.

C. Arduino Uno

Based on the ATmega328P microcontroller, Arduino Uno is an open-source microcontroller board. chip. It has a relatively small form factor and is designed for use in a wide range of electronic projects, including robotics, home automation, data logging, and IoT (Internet of Things) applications.

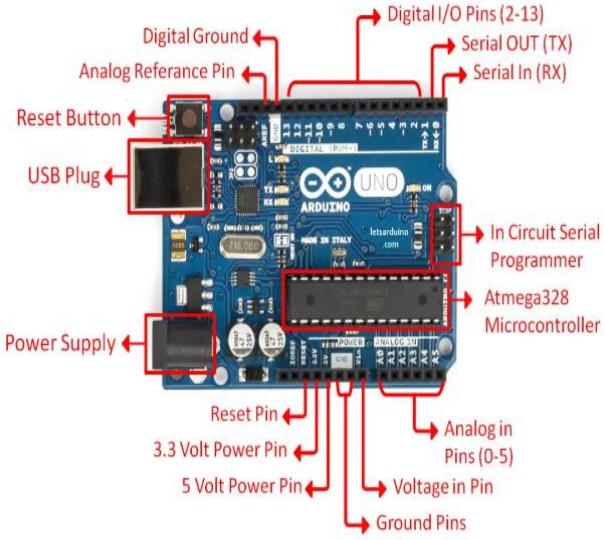

Figure 2: Pin diagram of Arduino Uno

The Arduino Uno board has a number of key features, including:

- 14 Digital Input/Output Pins: These pins can be used to interface with a variety of digital devices, such as sensors, LEDs, and switches.
- 6 Analog Input Pins: These pins can be used to measure analog signals, such as light or temperature.
- 16 MHz Clock Speed: This clock speed allows the microcontroller to run code quickly and efficiently.
- USB Connection: This connection allows the board to be programmed and powered through a computer.
- Power Jack: This jack allows the board to be powered by an external power supply.
- Reset Button: This button resets the board.

After connecting hardware components and getting data in Arduino ide next step is to build an ML model.

D. *Creating Machine Learning Model* :

1. Importing Libraries:

NumPy: It's a library from Python used to deal with arrays, and in addition, it provides features of matrices, linear algebra, etc.

Pandas: It's a Python library used to deal with datasets.

Matplotlib: This library is used for creating animated, static, and interactive visualizations in Python.

Sklearn: Python library with various effective methods for statistical modeling and machine learning.

Seaborn: It is a Python data visualization library built on top of the popular matplotlib library. Seaborn is used to creating more visually appealing and informative statistical graphics than Matplotlib

2. Data Cleaning:

At first, we imported the dataset from data.gov.in for crops requirement in different regions of India with the help of the pandas' library. We have a dataset with six columns, one is the Nitrogen column containing Nitrogen nutrient requirement and the others are Phosphorus, Potassium, Temperature, Humidity, and Rainfall column.

3. Selection and comparison of ML Algorithms:

Before selecting an algorithm, we need to first assess and compare different options, and then choose the one that best fits our specific dataset. Machine Learning is the best approach for providing a practical solution to the crop selection problem.

There are various ML algorithms that can be used for predicting crop. This study will consider the following algorithms for selection and accuracy comparison:

- *Naïve Bayes Classifier*: A popular machine learning algorithm for classification tasks is Naive Bayes. It depends on Bayes' hypothesis and accepts that the elements used to characterize the information are autonomous of one another. Given the input features, the probability of each class is calculated by the algorithm, and the class with the highest probability is chosen as the predicted class. It gives an accuracy of 98.01%.
- *Logistic Regression*: A classification algorithm called logistic regression is used to predict the probability of a target variable. There are only two classes that can be found in the target binary variable. At the point when strategic relapse is applied to our dataset it has given an accuracy of 95.22%.
- *Decision Tree:* The Decision Tree is a method of machine learning that can be used for classification and regression. The model appears as a tree, with every hub addressing a property or component, each branch addressing a choice rule, and each leaf hub meaning a class mark or anticipated esteem. The calculation separates the information into subsets in view of the information highlight values, determined to deliver subsets that are essentially as homogeneous as conceivable with respect to the objective variable. The procedure is recursive and goes on until a stopping point is reached. When applied to the dataset, it has produced an accuracy of 90%.
- *Random Forest*: Based on various data subsets, the algorithm creates decision trees and then makes predictions for each subset. The algorithm improves the system's solution by combining the results from a voting process. The output of Random Forest is more accurate because it trains the data using the bagging technique. RF has provided an accuracy of 99.09 percent for these data.

4. Model Building:

In this step, we will build a Random Forest model which is based on RF(Random Forest *Classifier*) as we got maximum accuracy and precision with RF. We selected dependent and independent features which will be required for crop prediction and classification and also to recommend fertilizers application using pre-existing values of required nutrients.

5. Evaluation:

We assessed the model on "accuracy" and "precision". We ended up with a precision of 96.36 percent and an accuracy of 99.09 percent.

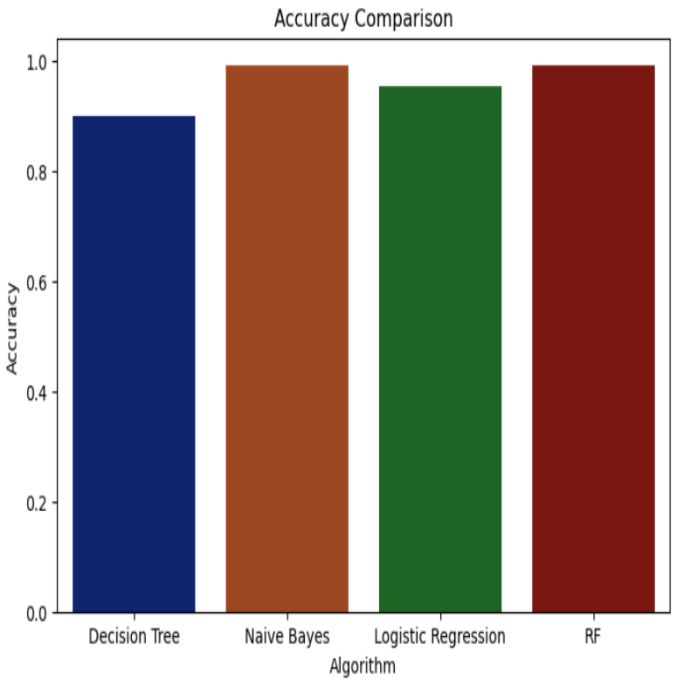

Figure 4: Accuracy comparison of ML models

E. Web application:

We have created a web application for accessing the machine learning model and taking input data from the user via the Soil moisture sensor and NPK sensor connected to Arduino. Application also uses weather API for temperature and humidity values.

This application consists of functions like crop prediction and fertilizer recommendation by machine learning techniques and data gathered from sensors.

3. Flowchart

We are using Arduino uno which will be getting sensor values to Arduino ide and storing it in datafile. After getting the data the web application process it, ML model will process the data and gives the required prediction for the crop and also recommends fertilizer if required.

The flow of the system is shown in following diagram

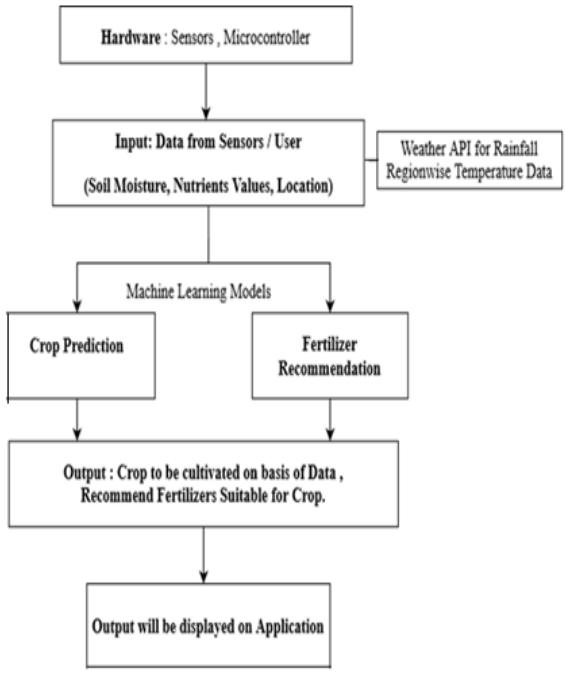

Figure 5: System flowchart

4. Results

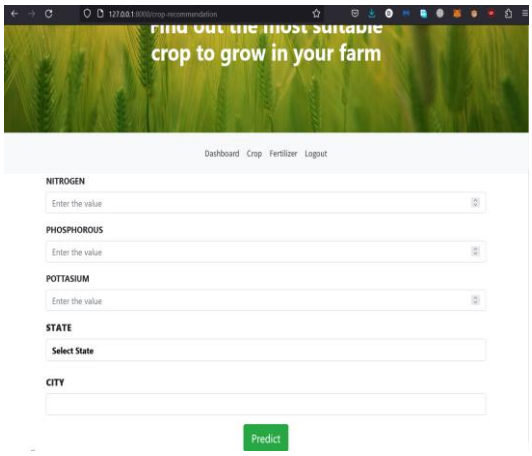

Figure 6: Crop prediction window

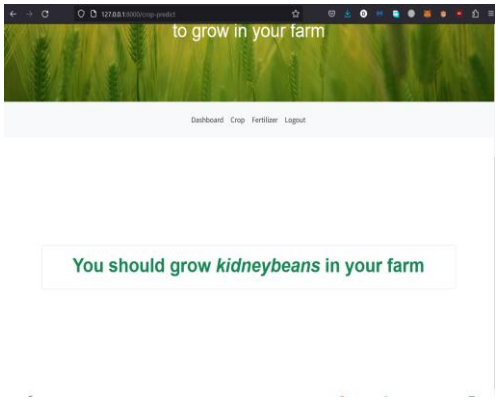

Figure 7: prediction result

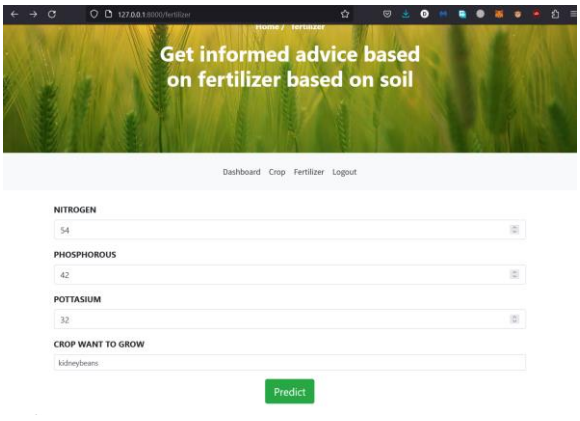

Figure 8: Fertilizer window

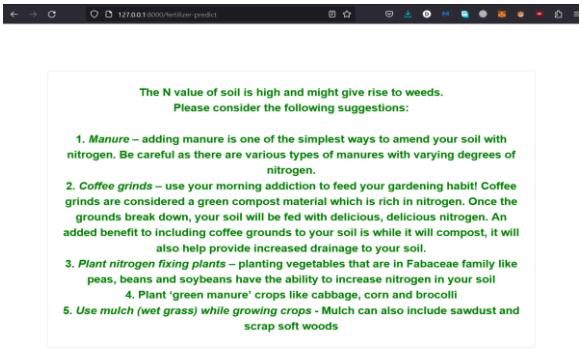

Figure 9: Fertilizer result

In the web application, we provided forms to accept data manually from the user and after pressing predict button, the ML model will process and shows the prediction result. The application requires login to access the prediction window which anyone can signup using the signup button.

5. Conclusion

This system focuses on the prediction of crop and fertilizers recommendations with the help of machine learning techniques and IoT sensors. Various machine-learning techniques were employed for the computation of precision. The Random Forest classifier was implemented to predict the harvest for a specific soil contains. A mechanism was established to anticipate crop yield based on prior data. The given System assists farmers in making choices regarding which crop to grow on their agricultural land. This system improves the

overall efficiency of farmland. By maximizing crop yield rates, this undertaking enhances the Indian economy.

References

[1] Amna Ikram, Waqar Aslam, Roza Hikmat Hama Aziz, Fazal Noor, Ghulam Ali Mallah , Sunnia Ikram, Muhammad Saeed Ahmad, Ako Muhammad Abdullah, and Insaf Ullah (2022). "Crop Yield Maximization Using an IoT-Based Smart Decision. In 2022 Journal of Sensor" Volume 2022, Article ID 2022923.

[2] Anakha Venugopal, Aparna S, Jinsu Mani, Rima Mathew, Prof. Vinu Williams. (2021). "Crop Yield Prediction using Machine Learning Algorithms". International Journal of Engineering Research & Technology (IJERT), NCREIS-2021 Conference Proceedings.

[3] Payal Gulati, Suman Kumar Jha. (2020). "Efficient Crop Yield Prediction in India using Machine Learning Techniques". In Proceedings of 2020 International Journal of Engineering Research & Technology (IJERT).

[4] Deepti Dighe, Harshada Joshi, Aishwarya Katkar, Sneha Patil, Prof. Shrikant Kokate. (2021). " Survey of Crop Recommendation Systems". In 2018 International Research Journal of Engineering and Technology (IRJET).

[5] Bhanu K N, Jasmine H J, Mahadevaswamy H S. (2020). "Machine learning Implementation in IoT based Intelligent System for Agriculture". International Conference for Emerging Technology (INCET), 2020, Belgaum, India.

[6] Jaimahaprabhu A, Praveen Kumar V, Gangadharan P S, Dr. B. Latha. (2019). "Cloud Analytics based Farming with Predictive Analytics using Artificial Intelligence". Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM).

[7] Aruvansh Nigam, Saksham Garg, Archit Agrawal "Crop Yield Prediction using ML Algorithms ",2019.

[8] Priya, P., Muthaiah, U., Balamurugan, M."Predicting Yield of the Crop Using Machine Learning Algorithm",2015.

[9] Tushar Hanwater, Madhav Vaidya, Ankush Sawarkar," Survey on Integrated platform for Development Rural Agriculture in India using ICT platform", International Journal of Computer Science and Information Technologies, Vol. 5 (2), 2014.

[10] Gouravmoy Bannerjee, Uditendu Sarkar, Swarup Das,& Indrajit Ghosh, "Artificial Intelligence in Agriculture: A Literature Survey," International Journal of Scientific Research in Computer Science Applications and Management Studies, Volume 7, Issue 3 (May 2018).