

# **PORTABLE CROP PREDICTION SYSTEM WITH MACHINE LEARNING ALGORITHM FOR AGRICULTURE**

Submitted in partial fulfilment of the requirements for the award of

Bachelor of Engineering degree in Mechatronics

By

**RISWANTH S (Reg. No. 40090024)**

**DEVARAGOTHAMAN Y A (Reg. No. 40090007)**



**DEPARTMENT OF MECHATRONICS**

**SCHOOL OF MECHANICAL ENGINEERING**

# **SATHYABAMA**

**INSTITUTE OF SCIENCE AND TECHNOLOGY**

**(DEEMED TO BE UNIVERSITY)**

**Accredited with grade "A++" by NAAC**

**JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI – 600119**

**APRIL - 2024**



**SATHYABAMA**

INSTITUTE OF SCIENCE AND TECHNOLOGY  
(DEEMED TO BE UNIVERSITY)

Accredited "A++" Grade by NAAC | 12B Status by UGC | Approved by AICTE

[www.sathyabama.ac.in](http://www.sathyabama.ac.in)

## DEPARTMENT OF MECHATRONICS

### BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of RISWANTH S (Reg. No. 40090024) and DEVARAGOTHAMAN Y A (Reg. No. 40090007) who carried out the project entitled **"PORTABLE CROP PREDICTION SYSTEM WITH MACHINE LEARNING ALGORITHM FOR AGRICULTURE"** under our supervision from December 2023 to April 2024.

Internal Guide

Dr. J.R. DEEPAK., M.E., Ph.D.,

Head of the Department

Dr. B. KANIMOZHI., M.E., PH.D.,

---

Submitted for Viva voce Examination held on 05.03.2024

Internal Examiner

External Examiner

## DECLARATION

We RISWANTH S (40090024) and DEVARAGOTHAMAN Y A (40090007) hereby declare that the Project Report entitled “**PORTABLE CROP PREDICTION SYSTEM WITH MACHINE LEARNING ALGORITHM FOR AGRICULTURE**” done by us under the guidance of **Dr. J. R. DEEPAK, M.E., Ph.D.**, is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering Degree in Mechatronics.

**DATE: 05.03.2024**

**PLACE: CHENNAI**

**1.**

**2.**

**SIGNATURE OF THE CANDIDATE**

## ACKNOWLEDGEMENT

We are pleased to acknowledge our sincere thanks to The Board of Management of **Sathyabama Institute of Science and Technology** for their encouragement in doing this project and for completing it successfully. We are grateful to them.

We convey our thanks to **Dr. B. KANIMOZHI., M.E., Ph.D.**, Head of the Department, Department of Mechatronics for providing us necessary support during the progressive reviews.

We would like to express our sincere and deep sense of gratitude to our Project Guide **Dr. J. R. DEEPAK., M.E., Ph.D.**, for his valuable guidance, suggestions and constant encouragement paved way for the successful completion of our project work.

We wish to express our thanks to all Teaching and Non-teaching staff members of the Department of Mechatronics who were helpful in many ways for the completion of the project.

We also specially thank our parents for their continuous support and cooperation throughout the end for the completion of the project.

## ABSTRACT

India, being among the top three cultivating countries, is renowned for its agricultural wealth, traditionally upheld through conventional farming methods, yet facing challenges due to recent environmental shifts affecting crop production. Recognizing this, a novel portable crop prediction system is introduced, utilizing sensors for real-time data collection on various environmental factors like temperature, humidity, pH, soil fertility, moisture, and rainfall. Integrated with machine learning and IoT technologies, the system amalgamates historical and real-time sensor data to forecast crucial agricultural parameters accurately. Leveraging algorithms such as MPL classifier, XG Boosting classifier, and bagging classifier, the system optimizes crop prediction, facilitated by an Arduino Mega and Zigbee modules, recommending suitable crops based on predictions. The system aims to enhance agricultural productivity by providing farmers with live sensor data display and actionable insights through a web-based platform, facilitated by Node MCU, enabling informed decision-making and adaptation to changing environmental conditions for optimal crop productivity. This promotes sustainable farming practices and efficiency, resonating with the environmentally conscious mindset of Gen Z farmers.

# TABLE OF CONTENTS

| CHAPTER<br>NO | TITLE                                                                                                          | PAGE<br>NO |
|---------------|----------------------------------------------------------------------------------------------------------------|------------|
|               | <b>ABSTRACT</b>                                                                                                | v          |
|               | <b>LIST OF FIGURES</b>                                                                                         | viii       |
|               | <b>LIST OF TABLES</b>                                                                                          | ix         |
|               | <b>LIST OF SYMBOLS AND ABBREVIATIONS</b>                                                                       | x          |
| <b>1</b>      | <b>INTRODUCTION</b>                                                                                            | 1          |
|               | 1.1 Agriculture                                                                                                | 1          |
|               | 1.2 Classification of farmers in India                                                                         | 1          |
|               | 1.2.1 Challenges faced by farmers in agriculture                                                               | 1          |
|               | 1.3 Various technologies to enhance crop yield                                                                 | 2          |
|               | 1.4 Machine learning                                                                                           | 3          |
|               | 1.5 Types of machine learning                                                                                  | 4          |
|               | 1.5.1 Supervised learning                                                                                      | 5          |
|               | 1.5.2 Unsupervised learning                                                                                    | 5          |
|               | 1.5.3 Reinforcement learning                                                                                   | 5          |
|               | 1.6 Types of machine learning algorithm                                                                        | 5          |
|               | 1.6.1 Regression                                                                                               | 5          |
|               | 1.6.2 Classification                                                                                           | 6          |
|               | 1.6.3 Clustering algorithms                                                                                    | 6          |
|               | 1.6.4 Q-Learning                                                                                               | 6          |
|               | 1.6.5 SARSA (State-Action-Reward-State-Action)                                                                 | 7          |
|               | 1.6.6 Markov Decision Process (MDP)                                                                            | 7          |
|               | 1.6.7 MLP Classifier (Multilayer Perceptron Classifier)                                                        | 7          |
|               | 1.6.8 XG Boost (Extreme Gradient Boosting)                                                                     | 8          |
|               | 1.6.9 Bagging Classifier                                                                                       | 8          |
|               | 1.7 Existing crop prediction systems                                                                           | 8          |
|               | 1.8 Challenges in the existing crop prediction system                                                          | 9          |
| <b>2</b>      | <b>LITERATURE SURVEY</b>                                                                                       | 10         |
|               | 2.1 Literature survey based on factors and sensors used for crop prediction                                    | 10         |
|               | 2.2 Literature review based on various machine learning algorithms and the parameters used for crop prediction | 11         |
| <b>3</b>      | <b>AIM AND SCOPE</b>                                                                                           | 14         |
|               | 3.1 Aim                                                                                                        | 14         |
|               | 3.2 Scope of the investigation                                                                                 | 14         |
| <b>4</b>      | <b>MATERIALS AND METHODS</b>                                                                                   | 16         |
|               | 4.1 Schematic diagram of the portable crop prediction system                                                   | 16         |
|               | 4.1.1 Updating data to IOT                                                                                     | 17         |

|          |                                                                                              |    |
|----------|----------------------------------------------------------------------------------------------|----|
|          | 4.1.2 Sending data to Python                                                                 | 17 |
|          | 4.2 Components used in the portable crop prediction system                                   | 18 |
|          | 4.3 Major components of the portable crop prediction system                                  | 18 |
|          | 4.4 Sensors used in the portable crop prediction system                                      | 22 |
|          | 4.5 Specifications of the sensors                                                            | 24 |
|          | 4.6 Power consumption analysis and battery selection for the portable crop prediction system | 25 |
|          | 4.7 Battery standby calculation                                                              | 26 |
|          | 4.8 Working of the portable crop prediction system                                           | 26 |
|          | 4.9 System Architecture of the portable crop prediction system                               | 27 |
|          | 4.10 Sequence diagram of the portable crop prediction system                                 | 29 |
|          | 4.11 Methodology of the portable crop prediction system                                      | 30 |
|          | 4.11.1 Data Pre-processing                                                                   | 30 |
|          | 4.11.2 Data Visualization                                                                    | 30 |
|          | 4.11.3 Algorithms used in the portable crop prediction system                                | 31 |
|          | 4.11.4 Evaluation methods of crop prediction algorithm                                       | 32 |
|          | 4.11.5 Confusion matrix testing for Algorithm accuracy                                       | 32 |
|          | 4.11.6 XG boost Classifier Algorithm Training ML Model                                       | 33 |
|          | 4.11.7 Dataset for training in XG Boost algorithm                                            | 37 |
|          | 4.12 Use case diagram of the portable crop prediction system                                 | 38 |
| <b>5</b> | <b>RESULTS AND DISCUSSION</b>                                                                | 40 |
|          | 5.1 Accuracy Test using confusion matrix                                                     | 40 |
|          | 5.2 Comparison of real time data with historical data                                        | 41 |
|          | 5.3 Real time data to check reliability                                                      | 44 |
|          | 5.4 Recommendations for the crops predicted from real time data based on soil nutrients      | 48 |
| <b>6</b> | <b>SUMMARY AND CONCLUSION</b>                                                                | 51 |
|          | 6.1 Summary                                                                                  | 51 |
|          | 6.2 Conclusion                                                                               | 51 |
|          | 6.3 Highlights of the portable crop prediction system                                        | 52 |
|          | <b>REFERENCES</b>                                                                            | 54 |
|          | <b>APPENDIX</b>                                                                              | 58 |

## LIST OF FIGURES

| FIGURE NO | TITLE OF FIGURE                                                                                                 | PAGE NO |
|-----------|-----------------------------------------------------------------------------------------------------------------|---------|
| 1.1       | Classification of farmers based on land size                                                                    | 1       |
| 1.2       | Various Machine learning algorithms                                                                             | 4       |
| 4.1       | Schematic diagram of the portable crop prediction system                                                        | 16      |
| 4.2       | Updating data to IOT                                                                                            | 17      |
| 4.3       | Sending data to python                                                                                          | 17      |
| 4.4       | Major components of the portable crop prediction system                                                         | 22      |
| 4.5       | Sensors used in the portable crop prediction system                                                             | 24      |
| 4.6       | 3D CAD model of the proposed portable crop prediction system                                                    | 27      |
| 4.7       | Overall process of the portable crop prediction system                                                          | 28      |
| 4.8       | Sequence diagram of the portable crop prediction system                                                         | 29      |
| 4.9       | Importing library files for XG Boost classifier                                                                 | 33      |
| 4.10      | Pre-processing using label encoder                                                                              | 33      |
| 4.11      | Training and testing of dataset                                                                                 | 34      |
| 4.12      | Distinguishing patterns and relationship within the dataset                                                     | 34      |
| 4.13      | Classification report of XG Boost classifier Algorithm                                                          | 35      |
| 4.14      | Confusion matrix test result of XG Boost Classifier                                                             | 36      |
| 4.15      | Hamming loss of XG Boost Classifier                                                                             | 37      |
| 4.16      | Graphical representation for the predicted result                                                               | 37      |
| 4.17      | Dataset for training                                                                                            | 38      |
| 4.18      | Use case diagram of the portable crop prediction system                                                         | 38      |
| 5.1       | Visualization of confusion matrices for ML Algorithms                                                           | 40      |
| 5.2       | Real time sensor data from different soil types displayed on the website of the portable crop prediction system | 45      |
| 5.3       | Real time sensor data displayed on LCD                                                                          | 45      |

## LIST OF TABLES

| TABLE NO | TITLE OF TABLE                                                                                         | PAGE NO |
|----------|--------------------------------------------------------------------------------------------------------|---------|
| 1.1      | Significant challenges influencing the growth and productivity of crops                                | 2       |
| 2.1      | Summary of sensors, parameters and machine learning algorithms used in various crop prediction systems | 10      |
| 2.2      | Summary of the various machine learning algorithms and the factors used                                | 12      |
| 4.1      | Technical specifications of Arduino mega 2560                                                          | 19      |
| 4.2      | Specification of sensors used in portable crop prediction system                                       | 25      |
| 4.3      | Power consumption of components                                                                        | 25      |
| 4.4      | Functionality of the confusion matrix                                                                  | 32      |
| 5.1      | Metrices comparison of algorithms used                                                                 | 41      |
| 5.2      | Sample real time sensor data collected for training the datasets                                       | 42      |
| 5.3      | Comparison between NPK values from sources and Observed NPK values from various soils                  | 42      |
| 5.4      | Real time sensor data of various factors that affect crop yield for different soil types               | 46      |
| 5.5      | Recommendations of soil nutrients                                                                      | 48      |

## LIST OF SYMBOLS AND ABBEREVIATIONS

|        |                                                             |
|--------|-------------------------------------------------------------|
| IoT    | Internet of Things                                          |
| NPK    | Nitrogen Phosphorus Potassium                               |
| AI     | Artificial Intelligence                                     |
| MCU    | Micro Controller Unit                                       |
| ML     | Machine Learning                                            |
| DBSCAN | Density-Based Spatial Clustering of Applications with Noise |
| SVM    | Support Vector Machines                                     |
| RNN    | Recurrent Neural Networks                                   |
| ANN    | Artificial Neural Network                                   |
| SVR    | Support Vector Regression                                   |
| KNN    | K-Nearest Neighbor                                          |
| SARSA  | State-Action-Reward-State-Action)                           |
| MDP    | Markov Decision Process                                     |
| MLP    | Multilayer Perceptron                                       |
| XGB    | Extreme Gradient Boosting                                   |
| LSTM   | Long Short-Term Memory                                      |
| CSM    | Crop Selection Method                                       |
| SRAM   | Static Random Access Memory                                 |
| LCD    | Liquid Crystal Display                                      |
| mg/kg  | milligrams per kilogram                                     |
| mA     | Milli Ampere                                                |

|        |                                                     |
|--------|-----------------------------------------------------|
| Ah     | Ampere hours                                        |
| Kg/ha  | Kilogram per hectare                                |
| ppm    | parts per million                                   |
| v      | volt                                                |
| C      | Celsius                                             |
| RS     | Recommended Standard                                |
| PWM    | Pulse with Modulation                               |
| USB    | Universal Serial Bus                                |
| SRAM   | Static Random Access Memory                         |
| EEPROM | Electrically Erasable Programmable Read Only Memory |
| LED    | Light Emitting Diode                                |
| HVAC   | Heating, Ventilation and Air Conditioning           |
| CAD    | Computer Aided Design                               |
| Wi-Fi  | Wireless Fidelity                                   |
| AC     | Alternating Current                                 |
| DC     | Direct Current                                      |

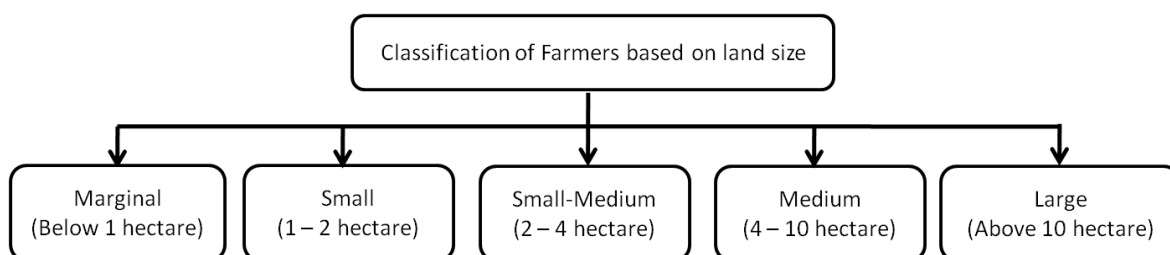
# CHAPTER 1

## INTRODUCTION

### 1.1 AGRICULTURE

Agriculture is crucial for ensuring food security and economic stability worldwide, particularly in countries like India where it is a primary occupation. It is supporting the economy of the country. In India agriculture plays a key role in setting the economy of the country. Nearly 50% of the Indian population completely depends on agriculture and has their primary occupation as farming.

### 1.2 CLASSIFICATION OF FARMERS IN INDIA



***Fig 1.1 Classification of farmers based on land size***

According to Indian agricultural, farmers are classifier into five categories based the size of the land they have which includes marginal, small, small-medium, medium and large. Moreover, a new era of young people called as Gen Z farmers have less amount of experience but a good amount of interest in agriculture not only as a profession but also to enhance agriculture all over the country as it is the wealth and key economy of India.

#### ***1.2.1 Challenges faced by farmers in Agriculture.***

Farmers face numerous challenges that affect crop growth and productivity. The challenges such as meteorological factors, soil factors, biotic factors and hazardous factors make agriculture very difficult for the farmers and demand them to develop innovative solutions to optimize and predict crop yields. Table 1.1 shows the

significant challenges influencing the growth and productivity of the crops, hence affecting Agriculture.

**Table 1.1 Significant challenges influencing the growth and productivity of crops**

| <b>Meteorological Factors</b>        | <b>Soil Factors</b> | <b>Biotic Factors</b>   | <b>Hazardous Factors</b> |
|--------------------------------------|---------------------|-------------------------|--------------------------|
| Sudden change in climatic conditions | Soil moisture       | Anthropogenic evolution | Drought                  |
| Rise in atmospheric temperature      | Soil fertility      | Pests and Weeds         | Flood                    |
| Amount of humidity and rainfall      | pH level            | -                       | Cyclones                 |

### 1.3 VARIOUS TECHNOLOGIES TO ENHANCE CROP YIELD

Traditionally, farmers relied on their experience to predict crop yields, but with the rapidly changing agricultural landscape, adaptation has become difficult. Machine learning has emerged as a game-changer in agriculture, using data-driven insights and predictive modelling to create automated crop yield prediction systems. This technology provides farmers with precise and timely information, helping them make informed decisions about crop selection, sowing/harvesting practices, soil fertility, and water management.

Machine learning (ML) has been widely applied in various fields, but its potential in agriculture remains largely untapped. ML offers a straightforward and efficient way to modernize farming practices for today's farmers. ML algorithms can analyze vast amounts of data related to soil quality, weather patterns, and historical data for analyzing. By doing so, they help farmers make better decisions about which crops to plant and how to manage resources like water and fertilizer. ML also helps predict future crop yields by real time sensor monitoring, enabling farmers to plan their harvests more effectively and manage their resources accordingly. This insight can be invaluable for farmers, helping them optimize their operations and increase their profitability.

The predictions generated by the model serve as valuable information for farmers, agricultural experts, and policymakers. They can use these predictions to make informed decisions about planting, irrigation, fertilization, and crop management. For example, if the model predicts a lower yield due to adverse weather conditions, farmers can take preventive measures to mitigate potential losses. In summary, integrating ML into agriculture has the potential to make farming more efficient, productive, and sustainable. As technology continues to advance, leveraging ML in agriculture could lead to significant improvements in crop yields and overall farm management.

These models integrate expert systems, environmental parameters, data mining algorithms, historical data, neural networks, and soil quality indicators to forecast yields. They offer farmers a comprehensive view of crop performance, empowering them to adjust strategies confidently.

#### 1.4 MACHINE LEARNING

Machine learning is the method of forecasting future results and outcomes by analyzing past existing data. It falls under the umbrella of artificial intelligence (AI) and enables computers to learn autonomously, without explicit programming. The essence of machine learning lies in crafting computer programs capable of adjusting to fresh data inputs. Python serves as the primary language for implementing basic machine learning algorithms. These algorithms undergo training using specialized techniques, where data is utilized to train the algorithm, enabling it to make predictions on new data sets. Machine learning is commonly categorized into three main types namely: supervised, unsupervised, and reinforcement learning.

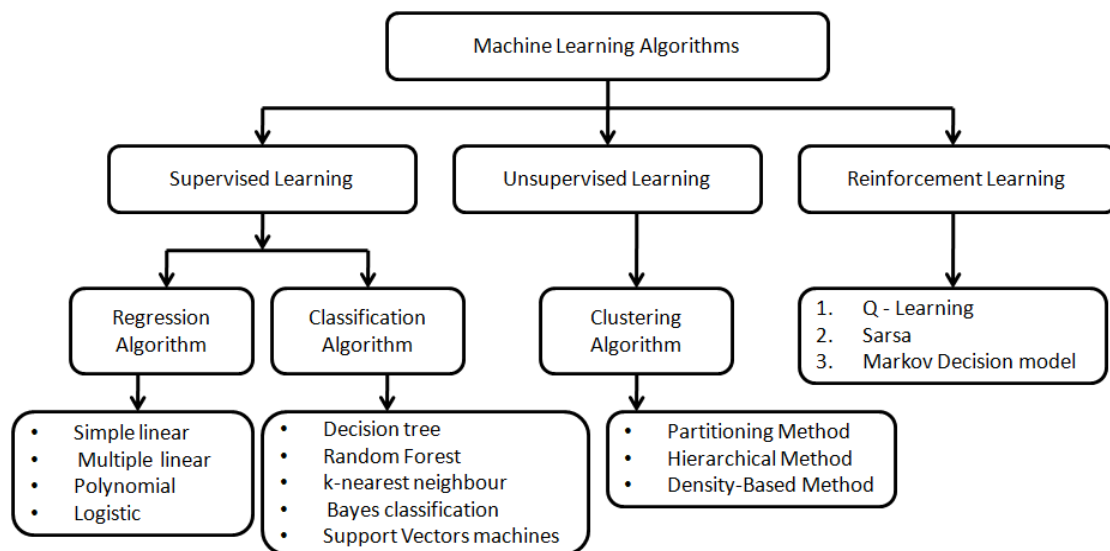
Supervised learning involves providing both input data and corresponding labels to the program for learning, with the labels being human assigned. In contrast, unsupervised learning operates on data lacking explicit labels, requiring the algorithm to determine data clustering autonomously. Reinforcement learning, on the other hand, interacts dynamically with its environment, receiving feedback (either positive value or negative value) to enhance its performance.

Researchers use numerous machine learning techniques in Python to uncover patterns that lead to actionable insights. These algorithms are roughly grouped into two categories based on how they learn from data for prediction purposes: supervised and unsupervised learning.

Classification is an important part of machine learning that involves predicting the class of supplied data pieces, which are sometimes referred to as targets, labels, or categories. Approximating a mapping function from input variables (X) to discrete output variables (y) is part of classification predictive modeling. It is a supervised learning approach in which computer software uses input data to classify new observations. The classification dataset might be binary (e.g., distinguishing between male and female, spam and non-spam emails) or multi-class.

## 1.5 TYPES OF MACHINE LEARNING

Machine learning involves different methods aimed at finding patterns in data, mainly divided into three types: supervised learning, unsupervised learning, and reinforcement learning.



**Fig 1.2 Various machine learning algorithms**

### ***1.5.1 Supervised Learning***

Supervised learning uses labeled data to train a model, with each input matching an output. This helps the model learn to make predictions or conclusions when presented with fresh, previously unknown data. It is used for tasks such as classification (labeling input data) and regression (predicting continuous values).

### ***1.5.2 Unsupervised Learning***

Unsupervised learning works with unlabeled data, requiring the model to recognize patterns without explicit supervision. Clustering methods combine comparable data points, whereas dimensionality reduction approaches simplify big datasets by lowering characteristics while maintaining critical information.

### ***1.5.3 Reinforcement Learning***

Reinforcement learning educates a machine in how to make decisions by interacting with its surroundings. It generates feedback based on its behaviors, which helps it develop the best techniques for maximizing rewards. This is beneficial in a variety of applications, including gaming, robotics, and autonomous systems.

These sorts of machine learning promote progress in a variety of industries, including healthcare, finance, autonomous cars, and recommendation systems. Practitioners as well as researchers use these methodologies to increase the possibilities for data analysis and decision-making.

## **1.6 TYPES OF MACHINE LEARNING ALGORITHM**

Under supervised learning, two main types of algorithms are commonly used: regression and classification.

### ***1.6.1 Regression***

When the output variable is a continuous value, regression methods are applied. The purpose of regression is to predict a numerical value using input information.

Linear regression, polynomial regression, and decision tree regression are among the most used regression techniques.

### ***1.6.2 Classification***

Classification algorithms are used when the output variable is a category or label. The goal of classification is to classify input data into predefined categories or classes. Common classification algorithms include logistic regression, decision trees, support vector machines (SVM), and naive Bayes. Logistic regression, for instance, predicts the probability that an input belongs to a certain class. Both regression and classification algorithms play crucial roles in supervised learning, enables to make predictions and decisions based on labelled data, these algorithms are widely used in various fields such as finance, healthcare, marketing, and beyond. Under unsupervised learning, one of the main types of algorithms is clustering.

### ***1.6.3 Clustering Algorithms***

Clustering algorithms are used to group similar data points together based on their characteristics or features. The primary goal of clustering is to identify patterns or structures within the data without predefined labels. Common clustering algorithms include k-means clustering, hierarchical clustering, and DBSCAN (Density-Based Spatial Clustering of Applications with Noise). In k-means clustering, the algorithm partitions the data into a predetermined number of clusters, with each cluster represented by its centroid. Data points are assigned to the cluster whose centroid is closest to them. Hierarchical clustering builds a tree-like hierarchy of clusters by iteratively merging or splitting existing clusters based on their similarity. Under reinforcement learning, several key algorithms are commonly used.

### ***1.6.4 Q-Learning***

Q-learning is a model-free reinforcement learning algorithm used to learn the optimal action-selection policy for an agent in each environment. The algorithm learns to estimate the value of taking a particular action in a specific state-learning

iteratively updates the Q-values based on the rewards received and the quality of actions taken. It aims to maximize the cumulative reward over time by selecting actions with the highest Q-values.

#### **1.6.5 SARSA (State-Action-Reward-State-Action)**

SARSA is another model-free reinforcement learning algorithm that learns the optimal policy by directly interacting with the environment. Unlike Q-learning, SARSA updates its Q-values based on the current action taken and the next action chosen. SARSA follows an on-policy learning approach, where the policy being learned is the same as the policy used to select actions during training. It is well-suited for scenarios where the agent's actions directly influence the environment's state.

#### **1.6.6 Markov Decision Process (MDP)**

MDP is a mathematical framework used to model decision-making in situations where outcomes are partially random and partially under the control of the decision-maker. In MDP, an agent makes decisions based on the current state of the environment and transitions to new states based on the actions taken. The agent receives rewards based on the transitions between states and actions chosen. MDPs provide a formal way to model reinforcement learning problems and are the foundation for algorithms like Q-learning and SARSA. These reinforcement learning algorithms are widely used in various applications, including robotics, gaming, autonomous systems, and recommendation systems. They enable agents to learn optimal strategies for maximizing long-term rewards in complex and uncertain environments.

#### **1.6.7 MLP Classifier (Multilayer Perceptron Classifier)**

The MLP classifier is a type of neural network algorithm used for classification tasks. It consists of multiple layers of nodes, including an input layer, one or more hidden layers, and an output layer. Each node in one layer is connected to every node in the subsequent layer with weighted connections. The MLP classifier uses

backpropagation to adjust the weights of these connections during training in order to minimize the error between the predicted and actual outputs.

#### **1.6.8 XG Boost (*Extreme Gradient Boosting*)**

XG Boost is an optimized and scalable implementation of gradient boosting, a machine learning technique used for regression and classification tasks. Gradient boosting works by building a series of weak learners (usually decision trees) sequentially, with each tree learning from the errors of its predecessors. XG Boost enhances traditional gradient boosting by employing a variety of regularization techniques and parallel computing to improve speed and model performance.

#### **1.6.9 Bagging Classifier**

Bagging (Bootstrap Aggregating) is an ensemble learning technique used to reduce the variance of a machine learning algorithm. In bagging, multiple base learners (e.g., decision trees) are trained independently on random subsets of the training data, with replacement. The predictions from these base learners are then combined (e.g., by averaging for regression or voting for classification) to make the final prediction. Bagging helps improve the stability and accuracy of the model by reducing overfitting and capturing different patterns in the data.

Top of Form

### **1.7 EXISTING CROP PREDICTION SYSTEMS**

- Intelligent Crop Recommendation System Using Machine Learning, forecasts crop production and provides the end user with accurate recommendations for fertilizer ratios depending on soil factors and also propose crop prices using machine learning.
- IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture was created using a multi-class support vector machine algorithm, and when combined with the IoT SNA-CR device, it allows the farmer to effortlessly keep his field soil information in the cloud

using his own mobile device with minimal experience.

- Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop cultivation, the model estimates production for a certain area, and sets the system's threshold. The threshold is compared to expected environmental data, and crop productivity is regarded the most important factor in making recommendations. The suggestion is made by comparing the threshold to the expected statistics for each season's crop. The best match is the suggested crop for the season in the region. This method takes the season and a list of crops as input and returns the suggested crop from the list.

## 1.8 CHALLENGES IN THE EXISTING CROP PREDICTION SYSTEMS

- Some crop prediction systems are developed considering Soil factors only eliminating the climatic conditions and change of weathers.
- The algorithms used in the earlier crop prediction systems may lack in accuracy when compared to recent machine learning algorithms that have been developed.
- Training the models with limited number of datasets may mislead to wrong recommendation of crops at times.
- Mobile apps developed were sometimes complicated to the farmers to easily access data and understand the recommendations.

## CHAPTER 2

### LITERATURE SURVEY

#### 2.1 LITERATURE REVIEW BASED ON FACTORS AND SENSORS USED FOR CROP PREDICTION

Based on the recent works carried out on crop yield prediction, a detailed survey is carried out in various factors that affect crop yield and various sensors and their specification, working parameters and the machine learning algorithms have been summarized in the table 2.1. It is observed that the sensors used include DHT11, FC-28, MQ-2, pH sensor, LM-393 and MQ-135 which are used to measure parameters like temperature and humidity, soil moisture, amount of CO<sub>2</sub>, pH level, sunlight and air quality respectively.

**Table 2.1 Summary of sensors, parameters and machine learning algorithms used in various crop prediction systems**

| Sensors   | Parameter                 | Range             | Accuracy   | Operating Voltage | Machine Learning Algorithm                                                                                   | Ref. |
|-----------|---------------------------|-------------------|------------|-------------------|--------------------------------------------------------------------------------------------------------------|------|
| DHT11     | Temperature, Humidity     | 0-50°C,<br>20-80% | 2°C,<br>5% | 3.3 V to 5.5 V    | Random forest classification,<br>K-Nearest neighbor classification,<br>Support vector machine classification | [10] |
| FC-28     | Soil Moisture             | 3.5V to 5.5V      | 99%        | 3.3 V to 5.5 V    |                                                                                                              |      |
| MQ-2      | Amount of CO <sub>2</sub> | 300 to 1000 ppm   | 80%        | 5 V               |                                                                                                              |      |
| pH Sensor | pH level                  | 3 – 9 pH          | ± 0.5 pH   | 4 V to 5 V        |                                                                                                              |      |

|           |                       |                |         |                |                            |      |
|-----------|-----------------------|----------------|---------|----------------|----------------------------|------|
| DHT11     | Temperature, Humidity | 0-50°C, 20-80% | 2°C, 5% | 3.3 V to 5.5 V | Naves bayes classification | [15] |
| FC-28     | Soil Moisture         | 3.5V to 5.5V   | 99%     | 3.3 V to 5.5 V |                            |      |
| DHT11     | Temperature, Humidity | 0-50°C, 20-80% | 2°C, 5% | 3.3 V to 5.5 V | Support Vector Clustering  | [17] |
| FC-28     | Soil Moisture         | 3.5V to 5.5V   | 99%     | 3.3 V to 5.5 V |                            |      |
| DHT11     | Temperature, Humidity | 0-50°C, 20-80% | 2°C, 5% | 3.3 V to 5.5 V | Support Vector machine     | [24] |
| FC-28     | Soil Moisture         | 3.5V to 5.5V   | 99%     | 3.3 V to 5.5 V |                            |      |
| MQ-135    | Air Quality           | 10~1000ppm     | 80%     | 2.5V to 5.0V   |                            |      |
| LM-393    | Sunlight              |                |         |                |                            |      |
| DHT11     | Temperature, Humidity | 0-50°C, 20-80% | 2°C, 5% | 3.3 V to 5.5 V | -                          | [25] |
| FC-28     | Soil Moisture         | 3.5V to 5.5V   | 99%     | 3.3 V to 5.5 V |                            |      |
| pH Sensor | pH level              | 0 to 14.00 pH  | 0.1 pH  | 1.75 V         |                            |      |

## 2.2 LITERATURE REVIEW BASED ON VARIOUS MACHINE LEARNING ALGORITHMS AND FACTORS USED FOR CROP PREDICTION

Based on a detailed literature review on various machine learning algorithms used for crop prediction system, it has been concluded that a diverse range of

classification and regression machine learning algorithms is utilized. To specify, random forest classifier, K-Nearest neighbor classifier, support vector machine classifier, Naive Bayes classification, support vector clustering, XG Boost classifier, logistic regression, gradient boosting regression, random forest regression, decision tree regression, kernel ridge regression, lasso regression, Enet algorithm, crop selection method (CSM) algorithm, and C 4.5 decision tree algorithm are among the algorithms employed to forecast crop yield production. The factors considered for crop yield prediction include temperature, humidity, rainfall, soil moisture, pH level, CO<sub>2</sub> levels, air quality, sunlight, pesticides, crop type, sowing time, production, plantation days, predicted yield rate, production, country, state, district, area, season, potential evapotranspiration, cloud cover, and wet day frequency. The literature review in Table 2.2 summarizes the various machine learning algorithms and the factors previously employed for predicting crop yield production.

**Table 2.2 Summary of the various machine learning algorithms and the factors used**

| Algorithm                                                                                                           | Factors                                                          | Ref. |
|---------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------|------|
| Random Forest Classifier<br>K-Nearest Neighbor Classifier<br>XG Boost Classifier<br>Logistic Regression             | Temperature, Rainfall,<br>Production                             | [4]  |
| Q learning Reinforcement<br>learning                                                                                | -                                                                | [8]  |
| Random forest classification,<br>K-Nearest neighbor<br>classification,<br>Support Vector machine<br>classification, | Humidity, Temperature,<br>pH, CO <sub>2</sub> ,<br>Soil Moisture | [10] |
| Gradient Boosting Regression,<br>Random Forest Regression,                                                          | Temperature, Rain,<br>Pesticides, Country, Yield                 | [12] |

|                                                        |                                                                                              |      |
|--------------------------------------------------------|----------------------------------------------------------------------------------------------|------|
| Support Vector Machine,<br>Decision Tree Regression    |                                                                                              |      |
| Random forest,<br>Decision tree                        | Rainfall, Temperature                                                                        | [14] |
| Naves bayes classification                             | Temperature, Humidity,<br>Soil Moisture, Rainfall, pH                                        | [15] |
| Support vector clustering                              | Temperature, Humidity,<br>Soil moisture                                                      | [17] |
| Kernel Ridge,<br>Lasso,<br>Enetalgorithms - Regression | State, District, Area,<br>Season                                                             | [20] |
| CSM algorithm                                          | Crop type, Sowing time,<br>Plantation days,<br>Predicted yield rate                          | [21] |
| Support Vector Machine                                 | Temperature, Humidity,<br>Air quality, Sunlight,<br>Moisture                                 | [24] |
| SQL based instead of ml                                | Temperature, Humidity,<br>Soil moisture, pH                                                  | [25] |
| decision tree (C 4.5)                                  | Rainfall, Temperature,<br>Potential<br>Evapotranspiration, Cloud<br>cover, Wet day frequency | [31] |

From the literature review, it is observed that the major factors that have been considered to affect crop yield previously include temperature, humidity and soil moisture. Based on this literature survey, this research work concentrates on the environmental factors including temperature, humidity, rainfall and soil factors such as soil fertility, pH of the soil and soil moisture. These factors are the key factors that affect crop yield and the real time data of above factors are collected using sensors such as DS18B20, DHT11, FC-37, pH sensor, soil NPK sensor, capacitive soil moisture sensor.

## CHAPTER 3

### AIM & SCOPE

#### 3.1 AIM

- Develop an innovative portable crop prediction system leveraging machine learning, IoT technology with Arduino and Zigbee for data collection in the field.
- Enhance agricultural productivity by accurately predicting crucial environmental factors such as temperature, humidity, rainfall, pH level, and soil fertility.
- Establish a web-based platform with Node MCU for live sensor data display and crop recommendations, aiding farmers in making informed decisions.
- Integrate historical and real-time sensor data to create a comprehensive environmental profile of agricultural land.

#### 3.2 SCOPE OF INVESTIGATION

- The portable crop prediction system aims to revolutionize traditional farming methods by integrating advanced technology to enhance crop yield prediction and management. By leveraging a combination of microcontrollers, sensor arrays, and machine learning algorithms, it provides farmers with real-time insights into crucial environmental factors affecting crop growth.
- The core components of portable crop prediction system include a microcontroller paired with an array of sensors such as Temperature, Humidity, Rainfall, pH, and NPK (Nitrogen, Phosphorus, Potassium). These sensors, including the DHT11 for temperature and moisture, Rain Sensor for rainfall, pH Sensor for soil acidity, and NPK Sensor for soil fertility, collect real-time data vital for crop health.

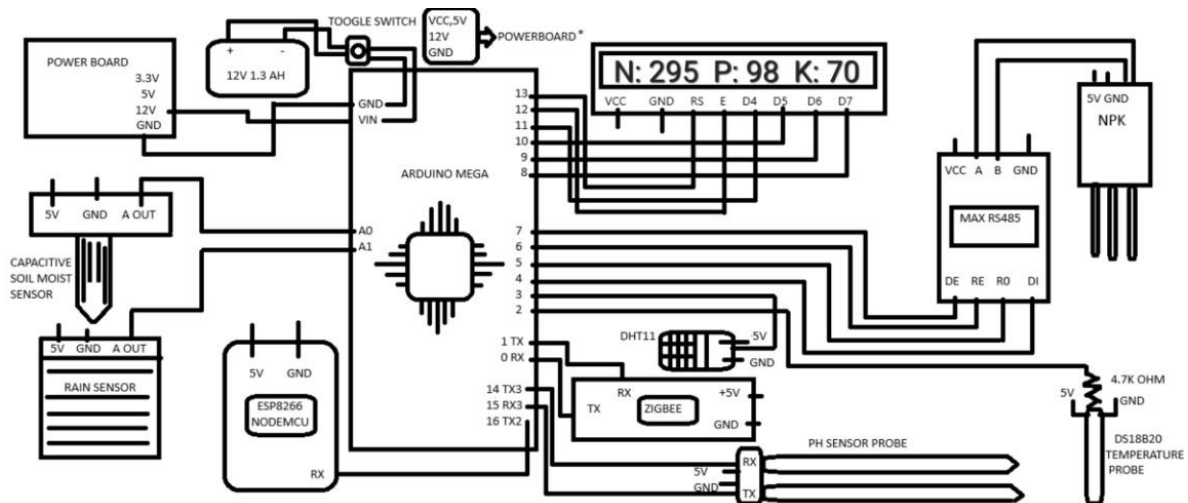
- To facilitate seamless data transfer and processing, Zigbee modules are employed. These modules enable the transmission of real-time sensor data to a laptop for further analysis and prediction using machine learning algorithms. By comparing current environmental conditions with historical datasets, the portable crop prediction system generates accurate forecasts of crop yield potential.
- Furthermore, to enhance user accessibility and convenience, Node MCU is incorporated to relay prediction results to a dedicated website. This website serves as a user-friendly platform where farmers can access live sensor data, receive crop recommendations tailored to their specific agricultural needs, and make informed decisions regarding crop management.
- In essence, this research work aims to empower farmers with actionable insights derived from advanced technology. By harnessing the power of smart agriculture and machine learning, it strives to optimize crop yield, minimize resource wastage, and promote sustainable farming practices for a more prosperous agricultural future.

## CHAPTER 4

### MATERIALS AND METHODS

#### 4.1 SCHEMATIC DIAGRAM OF THE PORTABLE CROP PREDICTION SYSTEM

Predicting the crop yield using historical data and comparing it with the real time data using IOT devices helps to increase the productivity of crops more effectively. Based on the recent research and developments in the field of agriculture, an innovative portable crop prediction system for agriculture is developed which collects real time data and predicts environmental factors such as atmospheric temperature, humidity, amount of rainfall, pH level and fertility of the soil. Figure 4.1 depicts the schematic diagram of the portable crop prediction system.



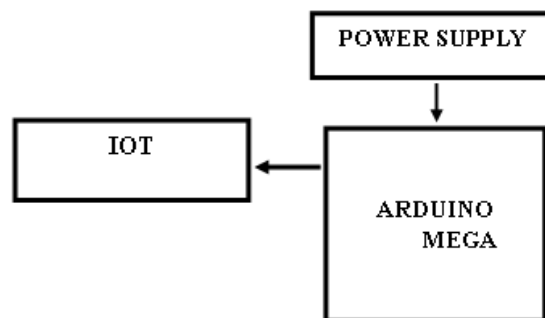
**Fig 4.1 Schematic diagram of the portable crop prediction system**

The portable crop prediction system is incorporated with a battery, Arduino Mega microcontroller and an array of sensors which include DS18B20 (Temperature Sensor), DHT11 (Humidity Sensor), pH Sensor, NPK Sensor (Soil Fertility Sensor), FC-37 (Rain Sensor), Capacitive Soil Moisture Sensor and Zigbee TX module as its input devices. These sensors contain probes which are inserted into the soil to collect the inputs and feed into the microcontroller. Using Zigbee transmitter and receiver modules, the data is transferred to the IOT system. The data is further transferred to the machine learning algorithm and the results are instantly relayed

to the dedicated website through Node MCU after the prediction, providing easy access to the users. By integrating the historical and real time sensor data, the crop prediction system creates a comprehensive environmental profile of the agricultural land using machine learning. This portable crop prediction system provides better sustainability and accurate predictions to the farmers there by increasing their productivity.

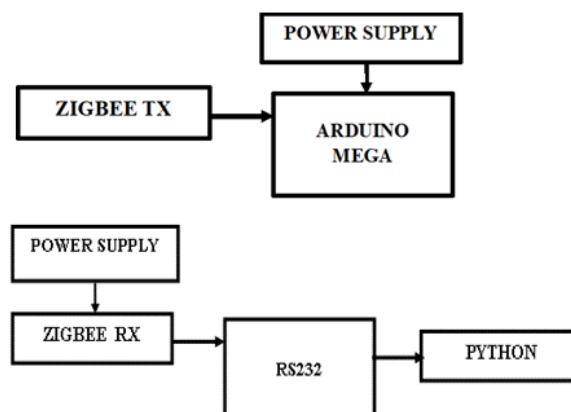
#### ***4.1.1 Updating data to IOT***

In this section, all the sensors' data will be updated through IOT webpage. Farmer can monitor all sensors data from anywhere and anytime also.



***Figure 4.2 Updating data to IOT***

#### ***4.1.2 Sending data to Python***



***Fig 4.3 Sending data to python***

In this section, all real time sensors data will be sent with the help of ZigBee communication module. There are two sections, transmitter's section will send the data. The receiver section will receive the data and will send it to python (machine learning algorithm) for crop prediction.

#### 4.2 COMPONENTS USED IN THE PORTABLE CROP PREDICTION SYSTEM

In the development of the crop prediction system, the following components have been integrated for real time sensing and transmitting the data to the algorithm.

These are the components used in the portable crop prediction system:

- Arduino Mega controller
- Power Supply
- 12V Battery
- LCD
- IoT Node MCU
- DHT11
- Toggle Switch
- pH Sensor
- Rain Sensor
- Capacitive Soil Moisture Sensor
- Zigbee TX-RX module
- DS18B20 Temperature sensor
- NPK Sensor
- RS485 Modbus Module

#### 4.3 MAJOR COMPONENTS OF THE PORTABLE CROP PREDICTION SYSTEM

- **Arduino Mega**

Table 4.1 provides a clear and concise overview of the technical specifications of the Arduino mega 2560. The Arduino Mega 2560 is a powerful microcontroller board designed for complex projects. It offers 54 digital input/output pins and 16 analog inputs, making it suitable for a wide range of applications, including 3D printers and robotics. With features like

PWM outputs, UARTs, a 16 MHz crystal oscillator, and USB connectivity, the Mega 2560 provides ample capabilities for project development. It's easy to use —simply connect it to a computer via USB or power it with an adapter or battery to get started. The Mega 2560 is highly versatile and offers plenty of space for sketching, making it a preferred choice for advanced projects. The AT mega2560's versatility and simplicity make it a valuable component across many industries and projects.

**Table 4.1 Technical specifications of Arduino mega 2560**

|                                |                                          |
|--------------------------------|------------------------------------------|
| Microcontroller                | AT mega 2560                             |
| Operating Voltage              | 5V                                       |
| Input Voltage<br>(recommended) | 7-12V                                    |
| Input Voltage (limit)          | 6-20V                                    |
| Digital I/O Pins               | 54 (of which 15 provide PWM output)      |
| Analog Input Pins              | 16                                       |
| DC Current per I/O Pin         | 20 mA                                    |
| DC Current for 3.3V Pin        | 50 mA                                    |
| Flash Memory                   | 256 KB of which 8 KB used by boot loader |
| SRAM                           | 8 KB                                     |
| EEPROM                         | 4 KB                                     |
| Clock Speed                    | 16 MHz                                   |
| LED_BUILTIN                    | 13                                       |
| Length                         | 101.52 mm                                |
| Width                          | 53.3 mm                                  |
| Weight                         | 37g                                      |

- **Power supply**

The power supply is pivotal for the smooth operation of the project, providing stable and regulated power. Utilizing a 0-12V/1 mA transformer, its primary connection links to the main supply through an on/off switch and a fuse,

safeguarding against overload and short circuits. In the power supply setup, the secondary part of the transformer is indispensable.

The adapter converts the incoming 12V AC voltage to 12V DC voltage, a critical step. Subsequently, capacitors filter the voltage, ensuring a consistent power flow. The final stage incorporates an IC 7805 regulator, maintaining a steady +5V output. This regulation is essential for the project's reliable and efficient operation, minimizing power fluctuations. The power supply includes 3.3V, 5V, and 12V output pins, catering to different requirements.

- ***LCD***

The LCD screen is an electronic display module widely used in various applications. A 16x2 LCD display, for instance, is a basic module commonly found in devices and circuits. Unlike seven segments and other LEDs, LCDs are preferred due to their cost-effectiveness, programmability, and ability to display special characters, animations, and custom content. The pin configuration of an LCD, crucial for its functionality, comprises several pins, each serving a distinct purpose. Among these pins are VSS and VDD, connecting to ground and the +5V power supply, respectively. The VO pin adjusts the display's contrast, while the RS pin selects between the command and data registers. RW, another essential pin, determines whether data is being read from or written to the LCD. The E pin enables the LCD for communication. Data pins D0 to D7 facilitate the transmission of data in 8-bit mode. Additionally, the A and K pins control the backlight's anode and cathode connections, respectively. Together, these pins enable effective communication with the LCD, facilitating the display of characters, numbers, and symbols.

- ***Node MCU***

The Node MCU is a versatile development board widely used in IoT projects and prototyping. It's based on the ESP8266 Wi-Fi module, offering built-in Wi-Fi connectivity that allows devices to connect to the internet and communicate with other devices or servers. The Node MCU board features GPIO pins for connecting various sensors, actuators, and peripherals,

making it suitable for a wide range of applications. Its compact size, ease of use, and compatibility with the Arduino IDE make it popular among hobbyists, students, and professionals alike for building connected projects and exploring the world of IoT.

- ***RS485 Modbus Module***

RS485 Modbus Module plays a crucial role in facilitating communication between different components, such as sensors and controllers. The RS485 Modbus Module enables reliable and efficient data transmission over long distances, making it suitable for industrial automation and control systems. It uses the Modbus protocol, a widely used communication protocol in the field of industrial automation, to exchange data between devices. By combining the RS485 Modbus Module with the NPK sensor, the system enables seamless communication and data exchange between the sensor and other components of the system.

- ***Zigbee TX-RX Module***

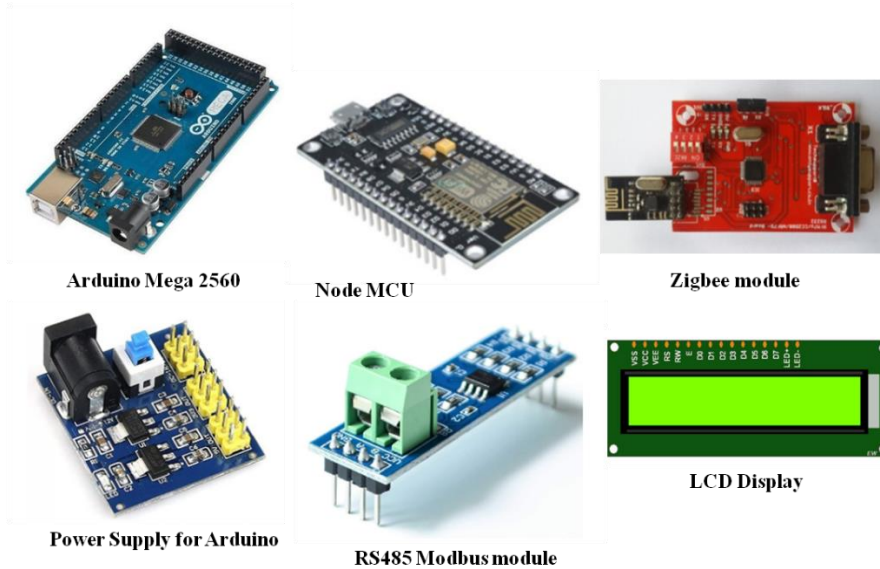
The Zigbee TX-RX module is a wireless communication device based on Zigbee technology. It consists of a transmitter (TX) and receiver (RX) pair that enable data transmission over long distances with low power consumption. The Zigbee module facilitates reliable and efficient communication between various devices in home automation, industrial monitoring, and IoT applications. Its mesh networking capability allows for seamless connectivity and data exchange among multiple nodes, making it suitable for building robust and scalable wireless networks.

- ***Toggle Switch***

A toggle switch is a simple, yet versatile component commonly used in electronic circuits and devices. It functions as a manual on/off switch, allowing users to control the flow of electricity by flipping a lever or toggle mechanism.

The switch is easy to operate and provides a reliable way to toggle between different states, making it ideal for a variety of applications such as lighting

controls, power supplies, and electronic gadgets. Its straightforward design and functionality are the reason for its wide application. Figure 4.4 shows the major components of the portable crop prediction system.



**Fig 4.4 Major components of the portable crop prediction system**

#### 4.4 SENSORS USED IN THE PORTABLE CROP PREDICTION SYSTEM

- **DHT 11**

The DHT11 is a basic yet effective digital sensor used to measure temperature and humidity levels in various applications. Its affordability and ease of use make it a popular choice for DIY projects and IoT setups. The sensor provides accurate readings of temperature and humidity, allowing users to monitor environmental conditions with precision. Its compact size and simple interface make it suitable for a wide range of projects, from weather stations to home automation systems.

- **Rain sensor**

The rain sensor, coupled with the LM393 comparator, offers a reliable solution for Arduino integration. The rain sensor detects the presence of moisture or raindrops, triggering the comparator LM393 to provide digital signals to the Arduino board. This integration allows for the creation of weather monitoring systems, automated irrigation systems, and rainfall

detectors. The simplicity and effectiveness of this setup make it ideal for various applications, including agriculture, gardening, and smart home projects.

- ***pH Sensor***

The pH sensor for Arduino is a valuable tool used to measure the acidity or alkalinity of a solution. It is equipped with TX and RX pins for communication with the Arduino board, allowing for easy integration into projects. Additionally, it often features a light diode that indicates the sensor's status or provides feedback during operation. This sensor enables users to accurately monitor pH levels in various liquids, making it useful for applications such as hydroponics, aquariums, and environmental monitoring.

- ***Capacitive Soil Moisture Sensor***

The capacitive soil moisture sensor is a vital component in agriculture and gardening projects. It measures the moisture content in the soil by utilizing capacitance changes. As the soil moisture level varies, the capacitance of the sensor changes accordingly, providing real-time data on soil moisture levels. This information helps in optimizing irrigation schedules, preventing overwatering or underwatering, and ensuring optimal conditions for plant growth. Its simplicity and accuracy make it an essential tool for maintaining healthy plants and maximizing crop yields.

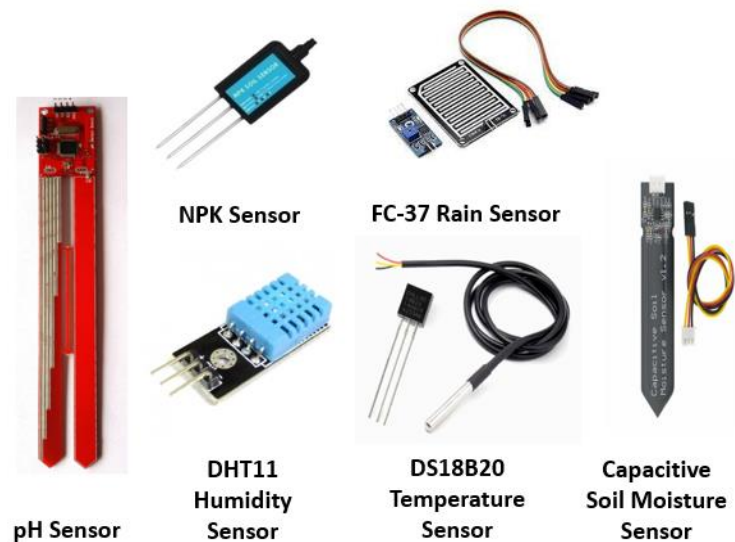
- ***DS18B20 Temperature Sensor***

The DS18B20 temperature sensor is a digital sensor used for precise temperature measurements in various applications. It provides accurate temperature readings with a resolution of up to 12 bits, making it suitable for temperature-sensitive systems and environments.

The sensor uses a one-wire communication protocol, enabling multiple sensors to be connected to a single microcontroller pin. Its small size, low cost, and high accuracy make it ideal for temperature monitoring in weather stations, HVAC systems, industrial processes, and scientific experiments.

- **NPK Sensor**

The NPK sensor is an essential tool for soil nutrient analysis and management in agriculture. It measures the levels of nitrogen (N), phosphorus (P), and potassium (K) – the primary nutrients required for plant growth and development. By providing real-time data on soil nutrient concentrations, the NPK sensor helps farmers optimize fertilizer usage, prevent nutrient deficiencies or excesses, and enhance crop productivity. The sensor's ability to monitor soil nutrient status enables precision farming practices, leading to improved soil health, sustainable agriculture, and higher crop yields. Figure 4.5 shows the sensors used in the portable crop prediction system.



***Fig 4.5 Sensors used in the portable crop prediction system***

#### 4.5 SPECIFICATIONS OF THE SENSORS

Table 4.2 depicts the overall main specifications of the sensors used in the proposed portable crop prediction system. It includes the operating voltage, sensing range, accuracy and the type of output signal of the various sensors used in the portable crop prediction system.

**Table 4.2 Specification of sensors used in portable crop prediction system**

| <b>Sensors</b>                            | <b>Operating voltage</b> | <b>Sensing range</b> | <b>Accuracy</b> | <b>Output signal</b> |
|-------------------------------------------|--------------------------|----------------------|-----------------|----------------------|
| DS18B20<br>(Temperature Sensor)           | 3 V – 5.5 V              | -55°C - 125°C        | ± 0.5°C         | Digital              |
| DHT11<br>(Humidity sensor)                | 3.3 V – 5.5 V            | 20 – 90 % RH         | ± 5 % RH        | Digital              |
| PH Sensor                                 | 4 V – 5 V                | 3 – 9 PH             | ± 0.5 PH        | Analog               |
| NPK Sensor<br>(Soil Fertility) with RS485 | 5 V – 30 V               | 0 – 1999 mg/Kg       | ± 2 %           | Digital              |
| FC-37<br>(Rain Sensor)                    | 3 V – 5 V                | 0 – 100 %            | 80 %            | Analog               |
| Capacitive Soil Moisture Sensor v2.0      | 3.3 V – 5.5 V            | 1.5 V – 5 V          | 99 %            | Analog               |

#### 4.6 POWER CONSUMPTION ANALYSIS AND BATTERY SELECTION FOR THE PORTABLE CROP PREDICTION SYSTEM

**Table 4.3 Power consumption of components**

| <b>Components</b>               | <b>Consumption in mA for an hour</b> |
|---------------------------------|--------------------------------------|
| Arduino Mega                    | 15                                   |
| LCD                             | 50                                   |
| IOT Module                      | 100                                  |
| DHT11                           | 1                                    |
| pH Sensor                       | 100                                  |
| Rain Sensor                     | 10                                   |
| Capacitive Soil Moisture Sensor | 10                                   |
| Zigbee TX-RX                    | 100                                  |
| DS18B20                         | 1                                    |
| NPK Sensor                      | 100                                  |

The proposed portable crop prediction system is integrated with a 12 V lead-acid battery. Since the battery can withstand for less time, the power consumption of each component plays an important role with the battery. Hence a power consumption analysis is made and tabulated in table 4.3.

#### 4.7 BATTERY STANDBY CALCULATION

If the standby current draw is 487 mA, the standby time is calculated using the same formula:

Taken:

- Battery Capacity: 1.3Ah
- Standby Current: 487 mA (convert to Amps by dividing by 1000, so it becomes 0.487A)

Using the formula:

Standby Time (hours) = Battery Capacity (Ah) / Standby Current (A)

Substituting the values:

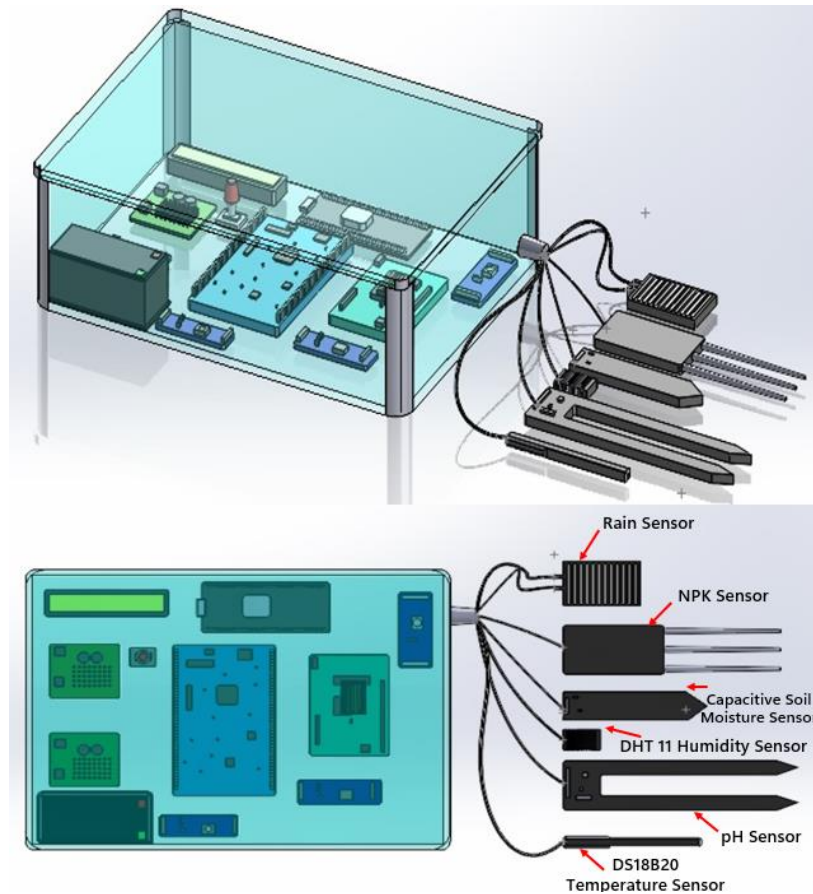
Standby Time (hours) = 1.3 Ah / 0.487 A  $\approx$  2.67 hours

So, with a 1.3Ah battery and a standby current draw of 487mA, the standby time would be approximately 2.67 hours.

#### 4.8 WORKING OF THE PORTABLE CROP PREDICTION SYSTEM

The crop prediction system is made up of several input devices, including a 12V lead-acid battery, an ATmega2560 microcontroller, a sensor array, a 16 x 2 LCD, an NRF24L01 (TX mode) Zigbee module, and an ESP-12E Node MCU Wi-Fi IoT module. These sensors, which include probes put into the soil, collect real-time data on temperature, humidity, pH level, soil fertility, rainfall, and soil wetness. The collected sensor data is then presented on an LCD attached to the system. When configured in TX mode (transmitter), the NRF24L01 provides sensor data to the output system, which then executes a machine learning algorithm. A toggle switch allows users to switch between the main AC power supply and the battery. Furthermore, the ESP-12E Node MCU Wi-Fi module enables the transmission of

live data to a dedicated website via an IoT platform, giving farmers easy access. Figure 4.6 depicts the 3D CAD model of the proposed portable crop prediction system.

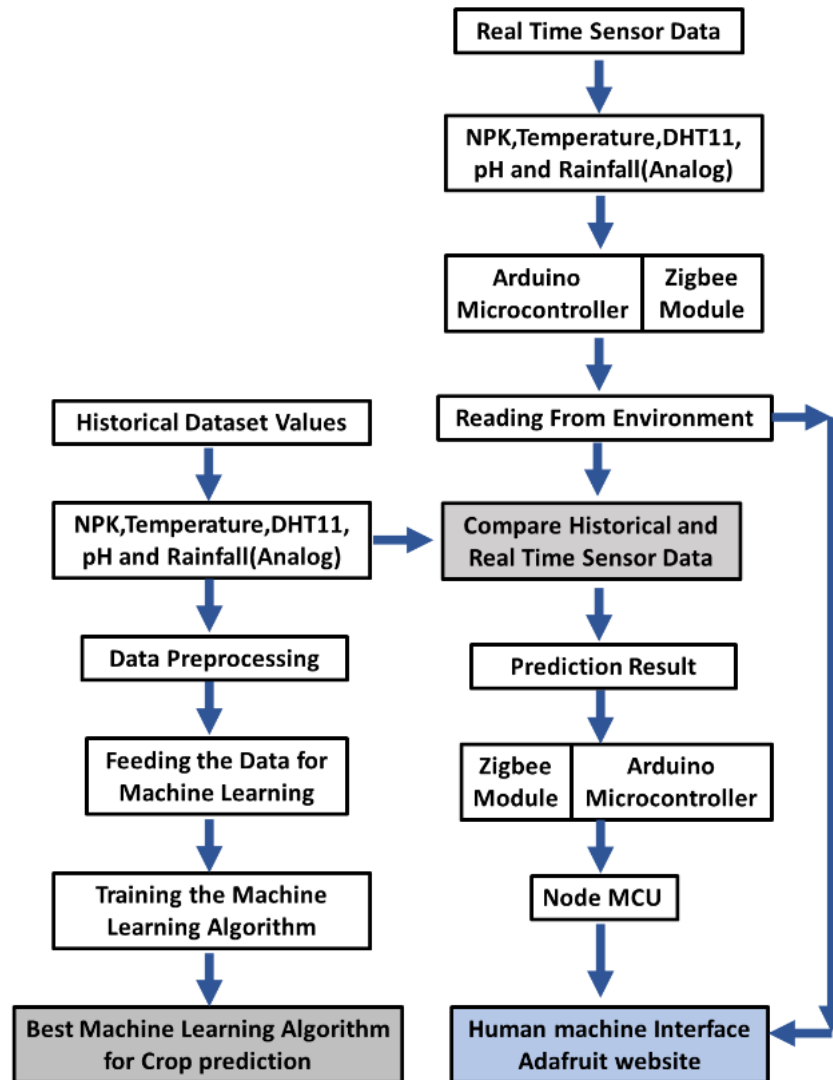


**Fig 4.6 3D CAD model of the proposed portable crop prediction system**

#### 4.9 SYSTEM ARCHITECTURE OF THE PORTABLE CROP PREDICTION SYSTEM

In the proposed system architecture, a comprehensive network is established for data acquisition, preprocessing, machine learning (ML) inference, and subsequent transmission. The system integrates multiple sensors, including NPK, temperature, DHT11, pH, rainfall (analog), and moisture sensors, which collectively capture crucial environmental parameters. These sensor data are initially collected by an Arduino microcontroller, which acts as a central hub for signal processing and data

aggregation. Utilizing Zigbee communication protocol, the Arduino seamlessly transmits the sensor readings to a receiving Zigbee module connected to a laptop. Figure 4.7 depicts the overall process of the portable crop prediction system.

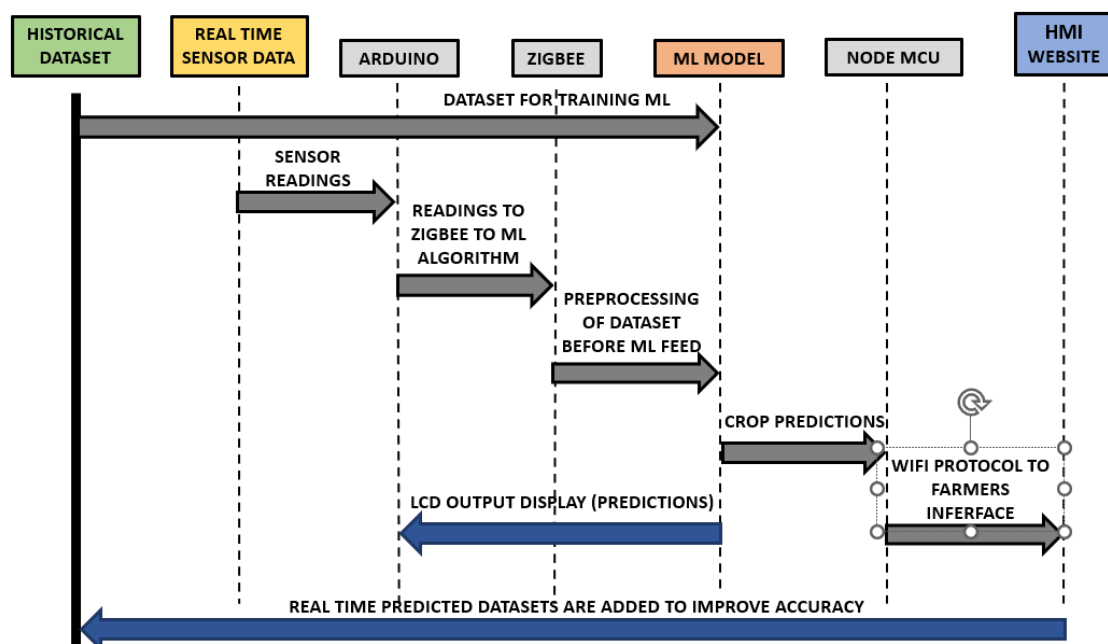


**Fig 4.7 Overall process of the portable crop prediction system**

Upon reception, the laptop processes the incoming data, preparing it for ML analysis. This preprocessing stage involves various techniques aimed at enhancing data quality and compatibility with the ML model. The ML model, pre-trained with a substantial dataset, undertakes the critical task of inference. It analyzes the incoming sensor data to derive insights, predict trends, or detect anomalies based on the learned patterns.

Once the ML algorithm completes its analysis, the resulting outputs are relayed back to the Arduino via the Zigbee network. This closed-loop communication loop ensures real-time feedback and control based on the ML-driven analysis. Additionally, a Node MCU device facilitates the transfer of processed data to an Adafruit platform through Wi-Fi connectivity. This integration enables remote monitoring and management of environmental conditions through the cloud based Adafruit interface.

#### 4.10 SEQUENCE DIAGRAM OF THE PORTABLE CROP PREDICTION SYSTEM



**Fig 4.8 Sequence diagram of the portable crop prediction system**

The system architecture comprises a sequential flow of operations aimed at real-time monitoring and analysis of environmental parameters. IoT sensors transmit data to an Arduino microcontroller, acting as the central hub for data aggregation. Through Zigbee communication, sensor readings reach a laptop for preprocessing before being fed into a pre-trained ML algorithm. Figure 4.8 depicts the sequence diagram of the portable crop prediction system.

- **Sensor Data Acquisition:** IoT sensors capture environmental values.

- **Data Transmission:** Arduino sends sensor data via Zigbee to the laptop.
- **Preprocessing:** The laptop preprocesses data to enhance compatibility with ML algorithms.
- **Machine Learning Analysis:** Pre-trained ML model analyzes data for insights and predictions.
- **Feedback Transmission:** ML outputs are sent back to the Arduino through Zigbee.
- **Wi-Fi Transmission:** Data is transferred from Arduino to NodeMCU and then to Adafruit via Wi-Fi.

By seamlessly integrating IoT technologies, data processing algorithms, and wireless communication protocols, the system provides a robust framework for precision agriculture and environmental monitoring.

#### 4.11 METHODOLOGY OF THE PORTABLE CROP PREDICTION SYSTEM

##### ***4.11.1 Data Preprocessing***

Data preprocessing refers to the initial steps and procedures performed on raw data before it is used for analysis or modeling. It involves cleaning, transforming, and organizing the data to make it suitable and efficient for further analysis. Data preprocessing tasks may include handling missing values, removing duplicates, scaling features, encoding categorical variables, and dealing with outliers. The goal of data preprocessing is to ensure that the data is accurate, complete, and in a format that can be effectively utilized by machine learning algorithms or other analytical methods.

##### ***4.11.2 Data Visualization***

Data visualization is the graphical representation of data and information using visual elements such as charts, graphs, and maps. The primary purpose of data visualization is to communicate insights, patterns, and trends in the data in a clear and understandable manner. By visually representing complex datasets, data visualization helps users to interpret and understand the underlying patterns and relationships within the data more easily. Effective data visualization techniques

enable data analysts, scientists, and decision-makers to explore data, identify patterns, detect outliers, and communicate findings effectively to a wide range of audiences.

#### ***4.11.3 Algorithms used in the portable crop prediction system***

Classification algorithms are a fundamental component of machine learning used to categorize data into predefined classes or labels. These algorithms play a crucial role in various fields, including agriculture, healthcare, finance, and more. In this research work, three distinct classification algorithms are explored namely: Multilayer Perceptron (MLP) Classifier, XG Boost (XGB) Classifier, and Bagging Classifier.

- ***Multilayer Perceptron (MLP) Classifier Algorithm***

The Multilayer Perceptron classifier belongs to the category of artificial neural networks (ANNs). It consists of multiple layers of nodes, each connected to the next layer. MLP classifiers are versatile and capable of learning complex patterns in data.

- ***Bagging Classifier Algorithm***

The Bagging Classifier is a type of ensemble learning method that combines multiple base classifiers trained on different subsets of the training data. It aggregates their predictions to produce a final output. Bagging is particularly effective in reducing overfitting and improving the stability of the model.

These classification algorithms utilize different underlying principles and techniques to classify data accurately. While MLP classifiers operate based on the principles of neural networks, XGB classifiers leverage gradient boosting, and Bagging classifiers employ ensemble learning strategies.

- ***XG Boost (XGB) Classifier Algorithm***

XG Boost is an implementation of gradient boosting algorithms, which belongs to the ensemble learning technique. It combines predictions from multiple weak models to create a stronger model. XG Boost is known for its efficiency, speed, and high performance in various classification tasks.

#### **4.11.4 Evaluation methods of crop prediction algorithm**

For the proposed system, the performance of these classification algorithms has been evaluated based on their accuracy scores. All three algorithms have been tested using confusion matrix test for accuracy scores and hamming loss. The accuracy score reflects the ability of the classifier to correctly classify data points. After thorough analysis, the XG Boost classifier is selected due to its superior accuracy score of 99.32%, making it the optimal choice for the portable crop prediction system.

#### **4.11.5 Confusion matrix testing for Algorithm accuracy**

The confusion matrix is like a report card for a classification model. It helps to see how well the model is doing by comparing what it predicted with what actually happened. It is the comparison between the actual values and the predicted values, and it has four constraints which is depicted in the table 4.4

- True Positive (TP): These are the times when the model got it right, predicting "yes" when it was actually "yes."
- False Positive (FP): Here, the model got it wrong, predicting "yes" when it was actually "no."
- True Negative (TN): These are the times when the model got it right, predicting "no" when it was actually "no."
- False Negative (FN): Here, the model got it wrong, predicting "no" when it was actually "yes."

**Table 4.4 Functionality of the confusion matrix**

| Classes             | Predicted Positive (1) | Predicted Negative (0) |
|---------------------|------------------------|------------------------|
| Actual Positive (1) | True Positive (TP)     | False Negative (FN)    |
| Actual Negative (0) | False Positive (FP)    | True Negative (TN)     |

By looking at the confusion matrix, the place where the model is making mistakes is observed. This helps to figure out where it is necessary to improve and make the

model better at predicting the right outcomes. Table 4.4 depicts the functionality of the confusion matrix.

#### **4.11.6 XG boost Classifier Algorithm Training ML Model**

Initially, it imports essential libraries such as pandas, numpy, and matplotlib, along with XG Boost and others. The dataset "CROP.csv" is loaded into a pandas Data Frame for further processing and analysis.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

import warnings
warnings.filterwarnings('ignore')

df = pd.read_csv('CROP.csv')
df.head()
```

**Fig 4.9 Importing library files for XG Boost classifier**

Subsequently, the categorical label column is pre-processed using Label Encoder, facilitating its conversion into integer values for enhanced computational compatibility. The dataset is then segregated into input features (X) and the target variable (y) for subsequent model training.

```
df=df.dropna()

df.columns

Index(['N', 'P', 'K', 'temperature', 'humidity', 'ph', 'rainfall', 'label'], dtype='object')

from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()

var = ['label']

for i in var:
    df[i] = le.fit_transform(df[i]).astype(int)
```

**Fig 4.10 Pre-processing using label encoder**

```

x = df.drop(labels='label', axis=1)
y = df.loc[:, 'label']

from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.20, random_state=42, stratify=y)
print("NUMBER OF TRAIN DATASET : ", len(x_train))
print("NUMBER OF TEST DATASET : ", len(x_test))
print("TOTAL NUMBER OF DATASET : ", len(x_train)+len(x_test))

NUMBER OF TRAIN DATASET : 1760
NUMBER OF TEST DATASET : 440
TOTAL NUMBER OF DATASET : 2200

print("NUMBER OF TRAIN DATASET : ", len(y_train))
print("NUMBER OF TEST DATASET : ", len(y_test))
print("TOTAL NUMBER OF DATASET : ", len(y_train)+len(y_test))

NUMBER OF TRAIN DATASET : 1760
NUMBER OF TEST DATASET : 440
TOTAL NUMBER OF DATASET : 2200

from xgboost import XGBClassifier

XGB = XGBClassifier()
XGB.fit(x_train,y_train)

```

**Fig 4.11 Training and testing of dataset**

A pivotal step involves the division of the dataset into training and testing subsets using the `train_test_split` function, allocating 80% of the data for training and 20% for testing. The dimensions of both training and testing datasets are meticulously documented for reference and validation purposes. The XG Boost classifier is instantiated and trained using the `fit` method, leveraging the training data to distinguish patterns and relationships within the dataset.

```

XGBClassifier(base_score=None, booster=None, callbacks=None,
              colsample_bylevel=None, colsample_bynode=None,
              colsample_bytree=None, device=None, early_stopping_rounds=None,
              enable_categorical=False, eval_metric=None, feature_types=None,
              gamma=None, grow_policy=None, importance_type=None,
              interaction_constraints=None, learning_rate=None, max_bin=None,
              max_cat_threshold=None, max_cat_to_onehot=None,
              max_delta_step=None, max_depth=None, max_leaves=None,
              min_child_weight=None, missing=nan, monotone_constraints=None,
              multi_strategy=None, n_estimators=None, n_jobs=None,
              num_parallel_tree=None, objective='multi:softprob', ...)

```

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.  
On GitHub, the HTML representation is unable to render, please try loading this page with [nbviewer.org](https://nbviewer.org).

```

predicted = XGB.predict(x_test)

from sklearn.metrics import classification_report
cr = classification_report(y_test,predicted)
print('THE CLASSIFICATION REPORT OF XGB CLASSIFIER:\n\n',cr)

```

**Fig 4.12 Distinguishing patterns and relationship within the dataset**

```
from sklearn.metrics import classification_report
cr = classification_report(y_test,predicted)
print('THE CLASSIFICATION REPORT OF XGB CLASSIFIER:\n\n',cr)
```

THE CLASSIFICATION REPORT OF XGB CLASSIFIER:

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0            | 1.00      | 1.00   | 1.00     | 20      |
| 1            | 1.00      | 1.00   | 1.00     | 20      |
| 2            | 1.00      | 1.00   | 1.00     | 20      |
| 3            | 1.00      | 1.00   | 1.00     | 20      |
| 4            | 1.00      | 1.00   | 1.00     | 20      |
| 5            | 1.00      | 1.00   | 1.00     | 20      |
| 6            | 1.00      | 1.00   | 1.00     | 20      |
| 7            | 1.00      | 1.00   | 1.00     | 20      |
| 8            | 0.95      | 1.00   | 0.98     | 20      |
| 9            | 1.00      | 1.00   | 1.00     | 20      |
| 10           | 1.00      | 0.90   | 0.95     | 20      |
| 11           | 1.00      | 1.00   | 1.00     | 20      |
| 12           | 1.00      | 1.00   | 1.00     | 20      |
| 13           | 0.95      | 1.00   | 0.98     | 20      |
| 14           | 0.95      | 1.00   | 0.98     | 20      |
| 15           | 1.00      | 1.00   | 1.00     | 20      |
| 16           | 1.00      | 1.00   | 1.00     | 20      |
| 17           | 1.00      | 1.00   | 1.00     | 20      |
| 18           | 1.00      | 1.00   | 1.00     | 20      |
| 19           | 1.00      | 1.00   | 1.00     | 20      |
| 20           | 1.00      | 0.95   | 0.97     | 20      |
| 21           | 1.00      | 1.00   | 1.00     | 20      |
| accuracy     |           |        | 0.99     | 440     |
| macro avg    | 0.99      | 0.99   | 0.99     | 440     |
| weighted avg | 0.99      | 0.99   | 0.99     | 440     |

**Fig 4.13 Classification report of XG Boost classifier Algorithm**

Following training, predictions are generated for the testing data using the predict method. A comprehensive classification report is then compiled using the classification\_report function from sklearn, providing insights into precision, recall, F1-score, and support for each class. The classifier's performance is further evaluated through the computation of a confusion matrix using sklearn's confusion\_matrix function. This matrix illuminates the model's accuracy and misclassifications across different classes, offering valuable diagnostic information.

Cross-validation is conducted using cross\_val\_score to gauge the model's consistency and generalization across various data splits, ensuring robust performance under diverse scenarios.

```
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test,predicted)
print('THE CONFUSION MATRIX SCORE OF XGB CLASSIFIER:\n\n',cm)
```

THE CONFUSION MATRIX SCORE OF XGB CLASSIFIER:

```
[[20  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0 20  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0 20  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0 20  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 20  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0 20  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0 20  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0 20  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 20  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0 20  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 20  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0 20  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0 20  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 20  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 20  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 20  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 20  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 20]]
```

```
from sklearn.model_selection import cross_val_score
accuracy = cross_val_score(XGB, x, y, scoring='accuracy')
print('THE CROSS VALIDATION TEST RESULT OF ACCURACY :\n\n', accuracy*100)
```

THE CROSS VALIDATION TEST RESULT OF ACCURACY :

```
[99.54545455 98.86363636 99.54545455 99.54545455 98.86363636]
```

```
from sklearn.metrics import accuracy_score
a = accuracy_score(y_test,predicted)
print("THE ACCURACY SCORE OF XGB CLASSIFIER IS :",a*100)
```

THE ACCURACY SCORE OF XGB CLASSIFIER IS : 99.31818181818181

**Fig 4.14 Confusion matrix test result of XG Boost Classifier**

Additionally, accuracy scores and hamming loss metrics are calculated to quantitatively assess the model's overall efficacy and predictive accuracy.

Lastly, a visual representation of the predicted versus actual values for the initial 100 samples is generated using matplotlib, facilitating intuitive comparison and validation of the classifier's predictions against ground truth data.

```

from sklearn.metrics import hamming_loss
hl = hamming_loss(y_test,predicted)
print("THE HAMMING LOSS OF XGB CLASSIFIER IS :",hl*100)

THE HAMMING LOSS OF XGB CLASSIFIER IS : 0.6818181818181818

def plot_confusion_matrix(cm, title='THE CONFUSION MATRIX SCORE OF XGB CLASSIFIER\n\n', cmap=plt.cm.cool):
    plt.imshow(cm, interpolation='nearest', cmap=cmap)
    plt.title(title)
    plt.colorbar()

cm1=confusion_matrix(y_test, predicted)
print('THE CONFUSION MATRIX SCORE OF XGB CLASSIFIER:\n\n')
print(cm)
plot_confusion_matrix(cm)

```

***Fig 4.15 Hamming loss of XG Boost Classifier***

This comprehensive process elucidates the systematic approach employed in training, evaluating, and validating the XG Boost classifier for crop prediction based on environmental parameters.

```

In [19]: import matplotlib.pyplot as plt
df2 = pd.DataFrame()
df2["y_test"] = y_test
df2["predicted"] = predicted
df2.reset_index(inplace=True)
plt.figure(figsize=(20, 5))
plt.plot(df2["predicted"][:100], marker='x', linestyle='dashed', color='red')
plt.plot(df2["y_test"][:100], marker='o', linestyle='dashed', color='green')
plt.show()

```

***Fig 4.16 Graphical representation for the predicted result***

In summary, classification algorithms play a pivotal role in predicting crop yields and optimizing agricultural productivity. By leveraging advanced machine learning techniques such as XG Boost, this research work aims to provide farmers with accurate predictions and valuable insights for effective crop management.

#### ***4.11.7 Dataset for training in XG Boost algorithm***

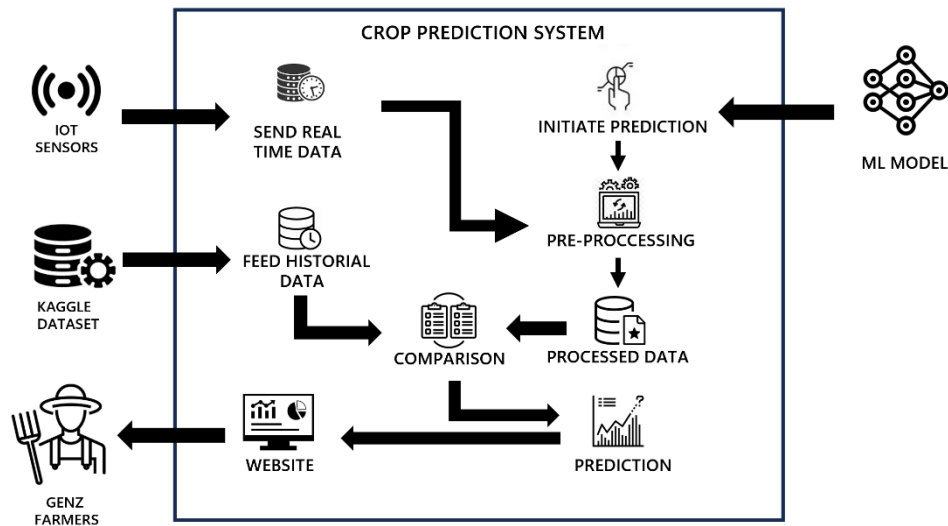
In this system, crucial environmental parameters including NPK (Nitrogen, Phosphorus, and Potassium), humidity, temperature, pH, and rainfall are sourced from Kaggle datasets for training purposes. These datasets serve as foundational components in the development and optimization of machine learning algorithms tailored for crop prediction and agricultural management. Through meticulous analysis and processing of these datasets, the system endeavors to harness the

insights embedded within the environmental data to enhance predictive accuracy and facilitate informed decision-making in agricultural practices.

| A                        | B                       | C                       | D                     | E            | F           | H     | I                                    |
|--------------------------|-------------------------|-------------------------|-----------------------|--------------|-------------|-------|--------------------------------------|
| N (mg kg <sup>-1</sup> ) | P(mg kg <sup>-1</sup> ) | K(mg kg <sup>-1</sup> ) | Temperature (celsius) | Humidity (%) | pH          | Label | rainfall(mm) daily avg ml feed value |
| 90                       | 42                      | 43                      | 20.87974371           | 82.00274423  | 6.502985292 | rice  | 2.254839291                          |
| 85                       | 58                      | 41                      | 21.77046169           | 80.31964408  | 7.038096361 | rice  | 2.51839486                           |
| 60                       | 55                      | 44                      | 23.00445915           | 82.3207629   | 7.840207144 | rice  | 2.932936084                          |
| 74                       | 35                      | 40                      | 26.49109635           | 80.15836264  | 6.980400905 | rice  | 2.698489269                          |
| 78                       | 42                      | 42                      | 20.13017482           | 81.60487287  | 7.628472891 | rice  | 2.919081561                          |
| 69                       | 37                      | 42                      | 23.05804872           | 83.37011772  | 7.073453503 | rice  | 2.789499998                          |
| 69                       | 55                      | 38                      | 22.70883798           | 82.63941394  | 5.70080568  | rice  | 3.014720671                          |
| 94                       | 53                      | 40                      | 20.27774362           | 82.89408619  | 5.718627178 | rice  | 2.688602166                          |
| 89                       | 54                      | 38                      | 24.51588066           | 83.5352163   | 6.685346424 | rice  | 2.560513732                          |
| 68                       | 58                      | 38                      | 23.22397386           | 83.03322691  | 6.336253525 | rice  | 2.457879953                          |
| 91                       | 53                      | 40                      | 26.52723513           | 81.41753846  | 5.386167788 | rice  | 2.940165219                          |
| 90                       | 46                      | 42                      | 23.97898217           | 81.45061596  | 7.50283396  | rice  | 2.778702596                          |
| 78                       | 58                      | 44                      | 26.80079604           | 80.88684822  | 5.108681786 | rice  | 3.160405074                          |
| 93                       | 56                      | 36                      | 24.01497622           | 82.05687182  | 6.98435366  | rice  | 2.058637099                          |
| 94                       | 50                      | 37                      | 25.66585205           | 80.66385045  | 6.94801983  | rice  | 2.32874412                           |
| 60                       | 48                      | 39                      | 24.28209415           | 80.30025587  | 7.042299069 | rice  | 2.567625941                          |
| 85                       | 38                      | 41                      | 21.58711777           | 82.7883708   | 6.249050656 | rice  | 3.073947177                          |
| 91                       | 35                      | 39                      | 23.79391957           | 80.41817957  | 6.970859754 | rice  | 2.29179095                           |
| 77                       | 38                      | 36                      | 21.8652524            | 80.1923008   | 5.953933276 | rice  | 2.495055743                          |
| 88                       | 35                      | 40                      | 23.57943626           | 83.58760316  | 5.85393208  | rice  | 3.236651798                          |
| 89                       | 45                      | 36                      | 21.32504158           | 80.47476396  | 6.442475375 | rice  | 2.061083036                          |

**Fig 4.17 Dataset for training**

#### 4.12 USE CASE DIAGRAM OF THE PORTABLE CROP PREDICTION SYSTEM



**Fig 4.18 Use case diagram of the portable crop prediction system**

The use case diagram illustrates how the crop prediction system assists modern farmers, particularly Gen Z farmers, in predicting suitable crops for their fields.

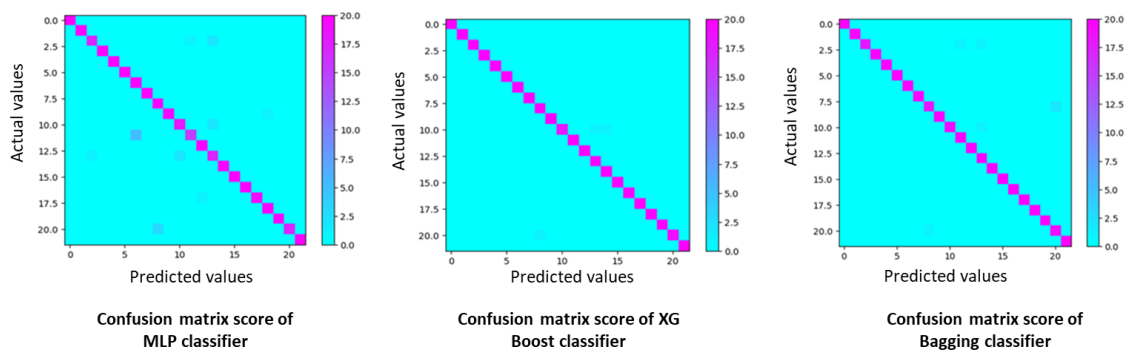
Figure 4.18 depicts the use case diagram of the portable crop prediction system. By analyzing environmental data over a minimum of 15 days, the system enables farmers to make informed decisions regarding crop production and sowing schedules. This iterative process empowers farmers to optimize their agricultural practices and enhance crop yield sustainability.

## CHAPTER 5

### RESULTS AND DISCUSSION

#### 5.1 ACCURACY TEST USING CONFUSION MATRIX

The historical data sourced from the Kaggle database is utilized to train the MLP Classifier, XG Boosting, and Bagging Classifier algorithms. The dataset is partitioned with 80% allocated for training and a random 20% for testing purposes. Machine learning program training is conducted based on the respective algorithms. The testing dataset undergoes evaluation with all three algorithms, and their confusion matrices are derived. Using these matrices, accuracy, precision, recall, F1 score, and weighted average metrics are computed and presented graphically for each algorithm. Observations reveal that the MLP Classifier, XG Boosting, and Bagging Classifier algorithms achieve accuracies of 95.90%, 99.31%, and 98.63%, respectively (Table 5.1). Figure 5.1 visually represents the confusion matrices for the MLP Classifier, XG Boosting, and Bagging Classifier algorithms. Hence, XG Boosting classifier algorithm is implemented to the system and the real time sensor data is processed, the most suitable crop is predicted which is displayed in the dedicated website. Further insight into the functionality of the confusion matrix is gained from Table 5.1.



**Figure 5.1 Visualization of confusion matrices for ML Algorithms**

The tabular classification provides a succinct overview of precision, F1 score, and accuracy metrics, enabling insightful comparisons across the models. XG Boost's

robust performance underscores its efficacy in predictive modelling and underscores its potential as a preferred choice for crop prediction systems.

**Table 5.1 Metrics comparison of algorithms used**

| MLP CLASSIFIER |        |          |         | XGB BOOST CLASSIFIER |        |          |         | BAGGING CLASSIFIER |        |          |         |
|----------------|--------|----------|---------|----------------------|--------|----------|---------|--------------------|--------|----------|---------|
| Precision      | Recall | F1-Score | Support | Precision            | Recall | F1-Score | Support | Precision          | Recall | F1-Score | Support |
| 1              | 1      | 1        | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 1              | 1      | 1        | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 0.94           | 0.85   | 0.89     | 20      | 1                    | 1      | 1        | 20      | 1                  | 0.9    | 0.95     | 20      |
| 1              | 1      | 1        | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 1              | 1      | 1        | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 1              | 1      | 1        | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 1              | 1      | 1        | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 0.8            | 1      | 0.89     | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 1              | 1      | 1        | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 0.87           | 1      | 0.93     | 20      | 0.95                 | 1      | 0.98     | 20      | 0.95               | 0.9    | 0.92     | 20      |
| 1              | 0.95   | 0.97     | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 0.9            | 0.9    | 0.9      | 20      | 1                    | 0.9    | 0.95     | 20      | 1                  | 0.95   | 0.97     | 20      |
| 0.94           | 0.75   | 0.83     | 20      | 1                    | 1      | 1        | 20      | 0.95               | 1      | 0.98     | 20      |
| 0.95           | 1      | 0.98     | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 0.81           | 0.85   | 0.83     | 20      | 0.95                 | 1      | 0.98     | 20      | 0.91               | 1      | 0.95     | 20      |
| 1              | 1      | 1        | 20      | 0.95                 | 1      | 0.98     | 20      | 1                  | 1      | 1        | 20      |
| 1              | 1      | 1        | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 1              | 1      | 1        | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 1              | 0.95   | 0.97     | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 0.95           | 1      | 0.98     | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 1              | 1      | 1        | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |
| 1              | 0.85   | 0.92     | 20      | 1                    | 0.95   | 0.97     | 20      | 0.9                | 0.95   | 0.93     | 20      |
| 1              | 1      | 1        | 20      | 1                    | 1      | 1        | 20      | 1                  | 1      | 1        | 20      |

## 5.2 COMPARISON OF REAL TIME DATA WITH HISTORICAL DATA

In this study, an effort is made to reduce reliance on Kaggle datasets by gathering and evaluating real-time data from the proposed system. This is done to assess the reliability of the dataset employed for program training. Initially 200 real time sample data have been collected from various districts that include, Chengalpattu, Chennai, Tiruvallur, and Kanchipuram. Table 5.2 shows the sample real time sensor data collected for training the datasets to the Machine Learning algorithm.

**Table 5.2 Sample real time sensor data collected for training the datasets**

| <b>N<br/>(kg/ha-1)</b> | <b>P<br/>(kg/ha-1)</b> | <b>K<br/>(kg/ha-1)</b> | <b>Temperature (Celsius)</b> | <b>Humidity (%)</b> | <b>pH</b> | <b>Rainfall (mm)</b> | <b>Crop</b>    |
|------------------------|------------------------|------------------------|------------------------------|---------------------|-----------|----------------------|----------------|
| 94                     | 50                     | 37                     | 25.66585                     | 80.66385            | 6.94802   | 2.328744             | rice           |
| 100                    | 48                     | 16                     | 25.71896                     | 67.22191            | 6.549902  | 0.827943             | maize          |
| 39                     | 29                     | 29                     | 26.50909                     | 94.48415            | 6.143663  | 2.220865             | coconut        |
| 104                    | 73                     | 46                     | 29.14009                     | 80.11902            | 6.282362  | 1.005016             | banana         |
| 30                     | 75                     | 25                     | 30.33277                     | 42.3525             | 6.446092  | 1.658888             | Pigeon<br>peas |
| 61                     | 68                     | 50                     | 35.21463                     | 91.49725            | 6.793245  | 2.700828             | papaya         |

Additionally, a comprehensive comparison is made between the recommended NPK (Nitrogen, Phosphorus, Potassium) values for different crops sourced from reputable websites and journals, and the observed ranges of these nutrients across four distinct soil types: beach soil, black soil, red soil, and alluvial soil. Table 5.3 shows the comparison between NPK values from sources and Observed NPK values from various soils. It highlights deficiencies and ideal levels crucial for optimal crop yield. This data aids in precisely addressing nutrient deficiencies, ensuring better agricultural outcomes. Examining the dataset reveals an immense amount of information on the complex interactions between crop choices, meteorological characteristics, and soil conditions.

**Table 5.3 Comparison between NPK values from sources and Observed NPK values from various soils**

| <b>Crop</b> | <b>Soil Nutrients (N, P, K)</b> | <b>Necessary Nutrients for crop yield (kg/ha-1)</b> | <b>Real time ranges from Beach soil (kg/ha-1)</b> | <b>Real time ranges from Black soil (kg/ha-1)</b> | <b>Real time ranges from red soil (kg/ha-1)</b> | <b>Real time ranges from Alluvial soil (kg/ha-1)</b> |
|-------------|---------------------------------|-----------------------------------------------------|---------------------------------------------------|---------------------------------------------------|-------------------------------------------------|------------------------------------------------------|
| Rice        | N                               | 150                                                 | 13-30                                             | 36-74                                             | 60-70                                           | 75-95                                                |
|             | P                               | 50                                                  | 14-34                                             | 50-65                                             | 40-55                                           | 45-80                                                |
|             | K                               | 50                                                  | 14-32                                             | 45-55                                             | 71-130                                          | 40-65                                                |
| Maize       | N                               | 150                                                 | 13-31                                             | 36-75                                             | 60-71                                           | 75-96                                                |
|             | P                               | 75                                                  | 14-30                                             | 50-66                                             | 40-56                                           | 45-81                                                |

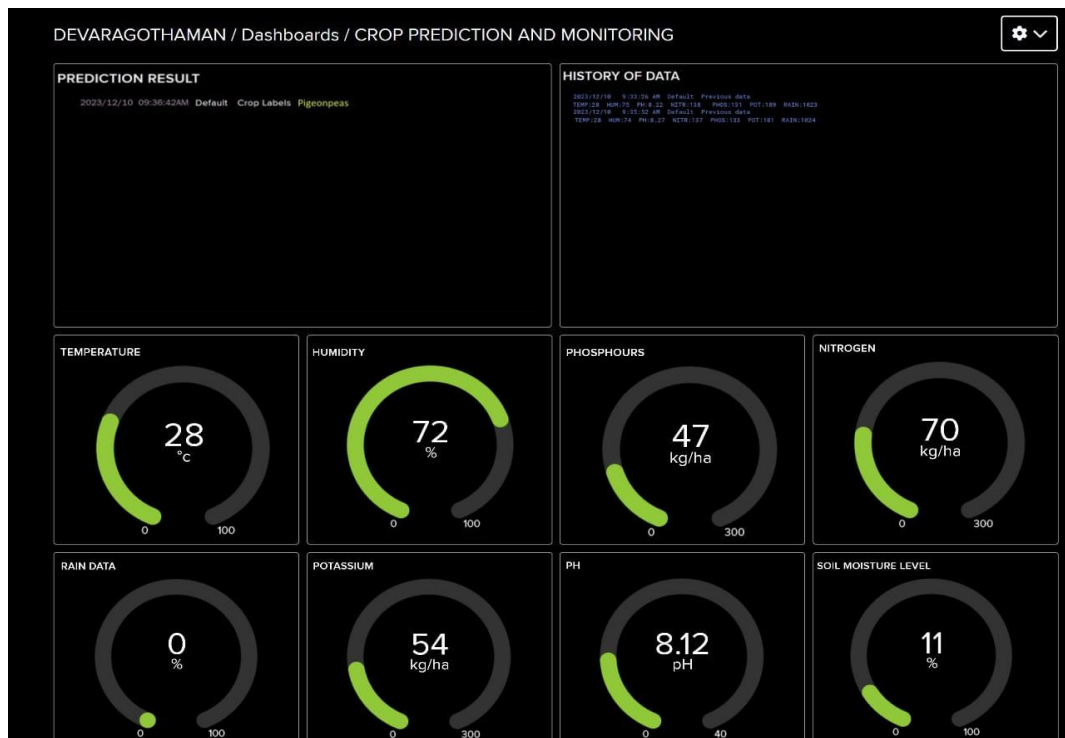
|             |   |     |       |       |        |       |
|-------------|---|-----|-------|-------|--------|-------|
|             | K | 75  | 14-28 | 45-56 | 71-131 | 40-66 |
| Chickpea    | N | 20  | 13-32 | 36-76 | 60-72  | 75-97 |
|             | P | 40  | 14-26 | 50-67 | 40-57  | 45-82 |
|             | K | 20  | 14-24 | 45-57 | 71-132 | 40-67 |
| Pigeon peas | N | 30  | 13-33 | 36-77 | 60-73  | 75-98 |
|             | P | 50  | 14-22 | 50-68 | 40-58  | 45-83 |
|             | K | 30  | 14-20 | 45-58 | 71-133 | 40-68 |
| Moth beans  | N | 40  | 13-30 | 36-78 | 60-70  | 75-99 |
|             | P | 40  | 14-34 | 50-69 | 40-55  | 45-84 |
|             | K | 40  | 14-32 | 45-59 | 71-130 | 40-69 |
| Black gram  | N | 25  | 13-31 | 36-79 | 60-71  | 75-95 |
|             | P | 50  | 14-30 | 50-70 | 40-56  | 45-80 |
|             | K | 25  | 14-28 | 45-60 | 71-131 | 40-65 |
| Pomegranate | N | 70  | 13-30 | 36-74 | 60-72  | 75-96 |
|             | P | 25  | 14-34 | 50-65 | 40-57  | 45-81 |
|             | K | 25  | 14-32 | 45-55 | 71-132 | 40-66 |
| Banana      | N | 140 | 13-31 | 36-75 | 60-73  | 75-97 |
|             | P | 50  | 14-30 | 50-66 | 40-58  | 45-82 |
|             | K | 200 | 14-28 | 45-56 | 71-133 | 40-67 |
| Mango       | N | 75  | 13-32 | 36-76 | 60-70  | 75-98 |
|             | P | 110 | 14-26 | 50-67 | 40-55  | 45-83 |
|             | K | 55  | 13-31 | 45-57 | 71-130 | 40-68 |
| Grapes      | N | 30  | 14-30 | 36-77 | 60-71  | 75-95 |
|             | P | 90  | 14-28 | 50-68 | 40-56  | 45-80 |
|             | K | 120 | 13-30 | 58-62 | 71-131 | 40-65 |
| Water melon | N | 35  | 14-34 | 36-78 | 60-78  | 75-96 |
|             | P | 25  | 14-32 | 50-69 | 40-57  | 45-81 |
|             | K | 25  | 13-31 | 45-59 | 71-132 | 40-66 |
| Muskmelon   | N | 40  | 14-30 | 36-79 | 60-73  | 75-97 |
|             | P | 60  | 14-28 | 50-70 | 40-58  | 45-82 |
|             | K | 30  | 13-32 | 45-60 | 71-133 | 40-67 |
| Apple       | N | 230 | 14-26 | 36-74 | 60-70  | 75-98 |
|             | P | 65  | 14-24 | 50-65 | 40-55  | 45-83 |

|         |   |     |       |       |        |       |
|---------|---|-----|-------|-------|--------|-------|
|         | K | 250 | 13-33 | 45-55 | 71-121 | 40-68 |
| Orange  | N | 30  | 13-30 | 36-75 | 60-71  | 75-95 |
|         | P | 12  | 14-34 | 50-66 | 40-56  | 45-80 |
|         | K | 12  | 14-32 | 45-56 | 71-122 | 40-65 |
| Papaya  | N | 94  | 13-31 | 36-76 | 60-72  | 75-96 |
|         | P | 72  | 14-30 | 50-67 | 40-57  | 45-81 |
|         | K | 75  | 14-28 | 45-57 | 71-123 | 40-66 |
| Coconut | N | 32  | 13-32 | 36-77 | 60-73  | 75-97 |
|         | P | 12  | 14-26 | 50-68 | 40-58  | 45-82 |
|         | K | 12  | 14-24 | 45-58 | 71-124 | 40-67 |

By contrasting the recommended nutrient values with the real-time soil ranges, it becomes evident which soil type closely aligns with the optimal nutrient requirements for each crop. For example, it is observed that the more amount of nutrients is required for better crop yield for crops like apple (230 kg/ha-1 of nitrogen, 65 kg/ha-1 of phosphorus, 250 kg/ha-1 of potassium) and banana (140 kg/ha-1 of nitrogen, 50 kg/ha-1 of phosphorus, 200 kg/ha-1 of potassium). This alignment signifies the potential for achieving the best crop yields under those specific soil conditions. However, where disparities exist between the recommended and observed nutrient levels, it highlights potential deficiencies that need to be addressed through targeted fertilization strategies. Thus, this comparative analysis serves as a valuable tool for farmers and agricultural practitioners to make informed decisions regarding soil management practices, ultimately leading to enhanced crop productivity and yield optimization.

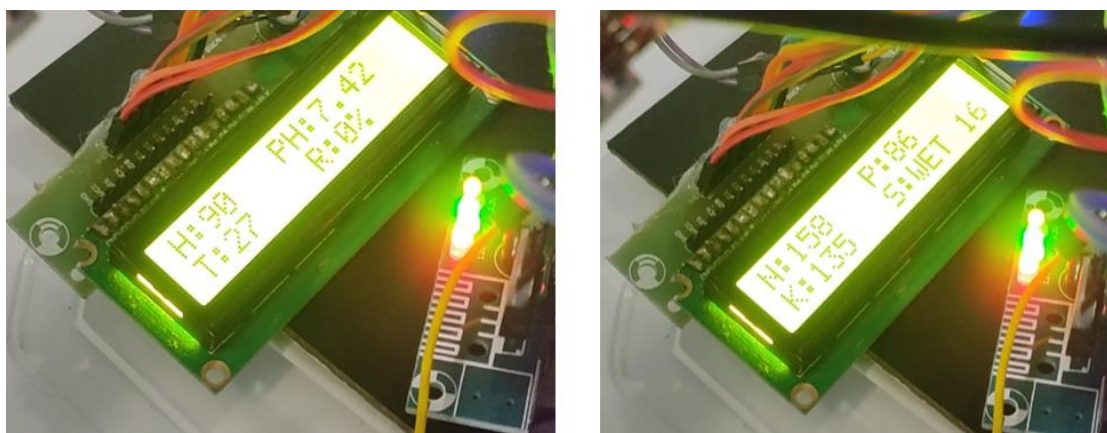
### 5.3 REAL TIME SENSOR DATA TO CHECK RELIABILITY

From previous studies, it is concluded that the major soil types that are available in India include alluvial soil, red soil, black soil beach soil. Hence, using the portable system, real time sensor data of the various factors including temperature, humidity, pH, nitrogen, phosphorus, potassium and rainfall are collected for consecutive weeks with a regular time interval during winter season from districts including Chennai, Chengalpattu, Kanchipuram and Tiruvallur.



**Figure 5.2 Real time sensor data from different soil types displayed on the website of the portable crop prediction system**

The real time sensor data obtained from various soil types are tabulated in table 5.4. Additionally, figure 5.2 shows the output result of real time sensor data displayed in the dedicated website of the portable crop prediction system and figure 5.3 shows the real time sensor data that is displayed on the LCD during the collection of real time data.



**Figure 5.3 Real time sensor data displayed on LCD**

**Table 5.4 Real time sensor data of various factors that affect crop yield for different soil types**

| Week 1        | Factors          |              |         |                    |                      |                     |                      |                     |
|---------------|------------------|--------------|---------|--------------------|----------------------|---------------------|----------------------|---------------------|
| Soil Type     | Temperature (°C) | Humidity (%) | pH (pH) | Nitrogen (kg ha-1) | Phosphorus (kg ha-1) | Potassium (kg ha-1) | Rain (Analog values) | Predicted Crop      |
| Beach Soil    | 28               | 95           | 6.28    | 24                 | 27                   | 26                  | 521                  | coconut             |
| Black Soil    | 28               | 92           | 8.32    | 70                 | 54                   | 47                  | 524                  | pigeon peas         |
| Alluvial soil | 27               | 81           | 6.98    | 90                 | 60                   | 62                  | 525                  | maize               |
| Red Soil      | 27               | 86           | 7.94    | 70                 | 51                   | 120                 | 119                  | banana              |
| Week 2        | Factors          |              |         |                    |                      |                     |                      |                     |
| Soil Type     | Temperature (°C) | Humidity (%) | pH (pH) | Nitrogen (kg ha-1) | Phosphorus (kg ha-1) | Potassium (kg ha-1) | Rain (Analog values) | Predicted Crop      |
| Beach Soil    | 26               | 94           | 6.14    | 27                 | 29                   | 29                  | 119                  | coconut             |
| Black Soil    | 28               | 72           | 8.12    | 68                 | 56                   | 47                  | 500                  | pigeon peas         |
| Alluvial soil | 25               | 81           | 7.68    | 90                 | 61                   | 63                  | 594                  | maize               |
| Red Soil      | 26               | 63           | 6.44    | 71                 | 52                   | 121                 | 327                  | banana              |
| Week 3        | Factors          |              |         |                    |                      |                     |                      |                     |
| Soil Type     | Temperature (°C) | Humidity (%) | pH (pH) | Nitrogen (kg ha-1) | Phosphorus (kg ha-1) | Potassium (kg ha-1) | Rain (Analog values) | Predicted Crop      |
| Beach Soil    | 27               | 65           | 6.42    | 27                 | 26                   | 28                  | 423                  | coconut             |
| Black Soil    | 29               | 81           | 6.84    | 67                 | 69                   | 47                  | 168                  | pigeon peas, papaya |
| Alluvial soil | 29               | 72           | 7.32    | 90                 | 61                   | 62                  | 524                  | maize               |
| Red Soil      | 29               | 68           | 8.04    | 70                 | 50                   | 121                 | 410                  | banana              |

| Week 4        | Factors          |              |         |                                 |                                   |                                  |                      |                |
|---------------|------------------|--------------|---------|---------------------------------|-----------------------------------|----------------------------------|----------------------|----------------|
| Soil Type     | Temperature (°C) | Humidity (%) | pH (pH) | Nitrogen (kg ha <sup>-1</sup> ) | Phosphorus (kg ha <sup>-1</sup> ) | Potassium (kg ha <sup>-1</sup> ) | Rain (Analog values) | Predicted Crop |
| Beach Soil    | 26               | 95           | 6.35    | 27                              | 27                                | 26                               | 468                  | coconut        |
| Black Soil    | 27               | 77           | 8.28    | 68                              | 52                                | 47                               | 231                  | pigeon peas    |
| Alluvial soil | 27               | 82           | 7.34    | 90                              | 59                                | 61                               | 924                  | maize          |
| Red Soil      | 28               | 81           | 8.14    | 70                              | 50                                | 120                              | 451                  | banana         |

From the table 5.4 it is observed that Beach soil, characterized by nitrogen ranging from 13-32 kg/ha, phosphorus from 14-26 kg/ha, and potassium from 14-24 kg/ha, exhibits relatively low nutrient levels with higher salt concentration, making it conducive for coconut cultivation. However, the soil slightly lacks 5-8 kg/ha of nitrogen for optimal growth. Addressing this deficiency through the addition of nitrogen can lead to perfect yield outcomes for coconut cultivation.

Pigeon peas, a recommended crop prediction by the system, showcases an ideal match in black soil, with its nutrient composition aligning closely with the recommended NPK values for pigeon peas cultivation. The soil's richness in nutrients ensures that pigeon peas are poised to achieve perfect yield outcomes. With the soil containing the recommended levels of nitrogen, phosphorus, and potassium for pigeon peas, black soil provides an optimal environment for maximizing crop productivity and ensuring successful harvests.

Alluvial soil exhibits a deficiency of 60 kg/ha of nitrogen, 14-17 kg/ha of phosphorus, and 13-14 kg/ha of potassium. However, the recommended crop prediction for this soil type is maize. By supplementing the soil with these required nutrients per hectare, maize cultivation can achieve optimal yield. With the addition of these fertilizers, alluvial soil holds the potential to provide a nurturing environment for maize crops, ensuring robust growth and maximizing yield outcomes.

Red soil, with its nutrient composition measured at 70 kg/ha of nitrogen, 50 kg/ha of phosphorus, and 120 kg/ha of potassium, is well-suited for banana cultivation, matching the recommended NPK values for this crop. While these soil conditions are conducive to banana growth, achieving average yield outcomes, for optimal results, an additional 70 kg/ha of nitrogen and 80 kg/ha of potassium are recommended to attain perfect yield. By addressing these specific nutrient requirements, banana cultivation in red soil can realize its full potential, ensuring enhanced productivity and quality harvests.

#### 5.4 RECOMMENDATIONS FOR THE CROPS PREDICTED FROM REAL TIME SENSOR DATA BASED ON SOIL NUTRIENTS

**Table 5.5 Recommendations of soil nutrients**

| <b>Week 1</b>        | <b>Factors</b>            |                             |                            | <b>Recommendations</b>                                                              |
|----------------------|---------------------------|-----------------------------|----------------------------|-------------------------------------------------------------------------------------|
| <b>Soil Type</b>     | <b>Nitrogen (kg ha-1)</b> | <b>Phosphorus (kg ha-1)</b> | <b>Potassium (kg ha-1)</b> |                                                                                     |
| <b>Beach Soil</b>    | 24                        | 27                          | 26                         | Soil slightly lacks in nitrogen, Addition of 8kg/ha N for Perfect yield             |
| <b>Black Soil</b>    | 70                        | 54                          | 47                         | Soil contains the recommended nutrients for Pigeon peas, so good yield is excepted  |
| <b>Alluvial soil</b> | 90                        | 60                          | 62                         | Addition of 60kg/ha N, 15kg/ha P & 13kg/ha K results in good Yield.                 |
| <b>Red Soil</b>      | 70                        | 51                          | 120                        | soil lacks nitrogen & potassium, Addition of 50kg/ha N & 30kg/ha for Perfect yield. |
| <b>Week 2</b>        | <b>Factors</b>            |                             |                            | <b>Recommendations</b>                                                              |
| <b>Soil Type</b>     | <b>Nitrogen (kg ha-1)</b> | <b>Phosphorus (kg ha-1)</b> | <b>Potassium (kg ha-1)</b> |                                                                                     |
| <b>Beach Soil</b>    | 27                        | 29                          | 29                         | Soil slightly lacks in nitrogen, Addition of 5kg/ha N for Perfect yield             |
| <b>Black Soil</b>    | 68                        | 56                          | 47                         | Soil contains the recommended nutrients for Pigeon peas, so good yield is excepted  |
| <b>Alluvial soil</b> | 90                        | 61                          | 63                         | Addition of 60kg/ha N, 14kg/ha P & 12kg/ha K results in good Yield.                 |

|                      |                           |                             |                            |                                                                                     |
|----------------------|---------------------------|-----------------------------|----------------------------|-------------------------------------------------------------------------------------|
| <b>Red Soil</b>      | 71                        | 52                          | 121                        | soil lacks nitrogen & potassium, Addition of 49kg/ha N & 29kg/ha for Perfect yield. |
| <b>Week 3</b>        | <b>Factors</b>            |                             |                            | <b>Recommendations</b>                                                              |
| <b>Soil Type</b>     | <b>Nitrogen (kg ha-1)</b> | <b>Phosphorus (kg ha-1)</b> | <b>Potassium (kg ha-1)</b> |                                                                                     |
| <b>Beach Soil</b>    | 27                        | 26                          | 28                         | Soil slightly lacks in nitrogen, Addition of 5kg/ha N for Perfect yield             |
| <b>Black Soil</b>    | 67                        | 69                          | 47                         | Soil contains the recommended nutrients for Pigeon peas, so good yield is excepted  |
| <b>Alluvial soil</b> | 90                        | 61                          | 62                         | Addition of 60kg/ha N, 14kg/ha P & 13kg/ha K results in good Yield.                 |
| <b>Red Soil</b>      | 70                        | 50                          | 121                        | soil lacks nitrogen & potassium, Addition of 50kg/ha N & 29kg/ha for Perfect yield. |
| <b>Week 4</b>        | <b>Factors</b>            |                             |                            | <b>Recommendations</b>                                                              |
| <b>Soil Type</b>     | <b>Nitrogen (kg ha-1)</b> | <b>Phosphorus (kg ha-1)</b> | <b>Potassium (kg ha-1)</b> |                                                                                     |
| <b>Beach Soil</b>    | 27                        | 27                          | 26                         | Soil slight lack in nitrogen, Addition of 5kg/ha N for Perfect yield                |
| <b>Black Soil</b>    | 68                        | 52                          | 47                         | Soil contains the recommended nutrients for Pigeon peas, so good yield is excepted  |
| <b>Alluvial soil</b> | 90                        | 59                          | 61                         | Addition of 60kg/ha N, 17kg/ha P & 14kg/ha K results in good Yield.                 |
| <b>Red Soil</b>      | 70                        | 50                          | 120                        | soil lacks nitrogen & potassium, Addition of 50kg/ha N & 30kg/ha for Perfect yield. |

The predictions of crop yield presented in table 5.4 are dependent not only on soil nutrient levels but also on other crucial factors such as temperature, humidity, pH, and rainfall. While these environmental variables play significant roles in determining crop growth and development, it's essential to highlight that the primary factors discussed in detail here are the NPK (Nitrogen, Phosphorus, Potassium) levels in the soil. These nutrients are fundamental to plant growth and are key determinants of crop yield. Therefore, while considering all environmental factors is

vital, the focus on NPK values provides critical insights into soil fertility and its direct impact on crop productivity. By understanding and addressing nutrient deficiencies, farmers can make informed decisions to optimize soil conditions and maximize crop yields. Table 5.5 suggest suitable recommendations of soil nutrients for the crops predicted from the real time data.

## CHAPTER 6

### SUMMARY AND CONCLUSION

#### 6.1 SUMMARY

From the provided text, it's evident that the predictions of crop yield rely not only on soil nutrient levels but also on various environmental factors such as temperature, humidity, pH, and rainfall. While these factors are crucial for crop growth and development, the focus primarily lies on the NPK (Nitrogen, Phosphorus, Potassium) levels in the soil due to their fundamental role in plant growth and crop yield determination.

By understanding and addressing nutrient deficiencies, farmers can optimize soil conditions and maximize crop yields. The crop prediction system presented employs a detailed approach by comparing recommended NPK values with real-time soil data across different soil types, effectively identifying deficiencies and ideal nutrient levels crucial for maximizing crop yields. This analysis helps farmers implement precise nutrient management strategies to address deficiencies and ensure better agricultural outcomes. Additionally, the portable crop prediction system, leveraging sensor technologies and machine learning algorithms, improves farming operations by providing real-time data collection on critical environmental elements, thus empowering farmers with comprehensive information about their agricultural area.

#### 6.2 CONCLUSION

In conclusion, the presented portable crop prediction system shows a detailed approach towards improving agricultural results. By comparing recommended NPK values with real-time soil data across various soil types, the system effectively identifies deficiencies and ideal nutrient levels crucial for maximizing crop yields. It has been noticed that some soils lack nutrients.

As a result, treating the soil with the optimal range of nutrients will result in a higher yield. According to the findings, the most appropriate crops for different soil types are coconut for beach soil, pigeon peas for black soil, maize for alluvial soil, and

banana for red soil. This thorough analysis provides valuable insights for farmers, helping them implement precise nutrient management strategies to address deficiencies and ensure better agricultural outcomes.

Moreover, the portable crop prediction system is a game changer in modern agriculture, leveraging sensor technologies and machine learning algorithms to improve farming operations. By including sensors for real-time data collecting on critical environmental elements such as temperature, humidity, pH, and soil fertility, the system provides farmers with full information on their agricultural area. The technology reliably forecasts crop production and proposes suitable crops using machine learning algorithms based on large datasets of historical and real-time data, allowing for more informed decision-making and farming practice improvement. Farmers benefit from the availability of live sensor data display and crop suggestions through a web-based platform, which provides accessibility and usability.

Furthermore, by enabling precise environmental monitoring and data-driven decision-making, the system not only improves crop output and quality, but also encourages sustainable farming practices and resource efficiency. The technology, which is specifically adapted to the environmentally conscious perspective of Gen Z farmers, promotes resource conservation through real-time optimization of fertilizer and water consumption, adding to its attractiveness and potential influence on modern agriculture. Overall, the portable crop prediction system is a significant step toward a more efficient, sustainable, and technologically advanced future for farming.

### 6.3 HIGHLIGHTS OF THE PORTABLE CROP PREDICTION SYSTEM

The key highlights of the portable crop prediction system are listed below:

- The portable crop prediction system integrates sensors for real-time data on environmental factors like temperature, humidity, pH, soil fertility, moisture, and rainfall.

- Sensors collect real-time data, which is processed using machine learning algorithms (XG Boost algorithm) which is trained using large datasets of historical data and real time data to predict crop yield and recommend suitable crops.
- Provides live sensor data display and crop recommendations via a web-based platform.
- Enhances decision-making by providing comprehensive environmental profiles of agricultural land, enabling informed choices and optimization of farming practices.
- Increases crop yield and quality through precise environmental monitoring and data-driven decision-making.
- The system provides real-time soil nutrient data, allowing Gen Z farmers to compare recommended nutrient values for predicted crops. This facilitates precise adjustment of soil nutrients, ensuring optimal conditions for maximum crop yield without the need for fertilizer recommendations. Additionally, it also reduces resource wastage by optimizing fertilizer and water usage based on real-time soil and weather conditions.
- Promotes sustainable farming practices and efficiency, resonating with the environmentally conscious mindset of Gen Z farmers.

## REFERENCES

- [1] Abhinav Sharma, Arpit Jain, Prateek Gupta and Vinay Chowdary, "Machine Learning Applications for Precision Agriculture: A Comprehensive Review", IEEE ACCESS, vol 9, pp.4843 - 4873, Dec.2020.
- [2] Agarwal S and Tarar S, "A Hybrid Approach for Crop Yield Prediction Using Machine Learning and Deep Learning Algorithms," J Phys Conf Ser, vol. 1714, no. 1, p. 012012, 2021.
- [3] Akshay Krishnan, Shashank Swarna, Balasubramanya H. S, "Robotics, IoT, and AI in the Automation of Agricultural Industry: A Review", in Bangalore Humanitarian Technology Conference (B-HTC), IEEE, 2020.
- [4] Aruvansh Nigam, Saksham Garg, Archit Agrawal, Parul Agrawal, "Crop Yield Prediction Using Machine Learning Algorithms", in Fifth International Conference on Image Information Processing (ICIIP), May.2019.
- [5] Chandrababha M, Rajesh Kumar Dhanaraj, "Machine learning based Pedantic Analysis of Predictive Algorithms in Crop Yield Management", in Fourth International Conference on Electronics, Communication and Aerospace Technology (ICECA-2020), Nov.2020.
- [6] Deepa N and Ganesan K, "Decision-making tool for crop selection for agriculture development," Neural Comput Appl, vol. 31, no. 4, pp. 1215–1225, 2019.
- [7] Deshmukh T, Rajawat A S, and Potgantwar A, "Machine Learning Technique for Crop Selection and Prediction of Crop Cultivation," in International Conference on Advanced Computing Technologies and Applications (ICACTA), IEEE, pp. 1–7, 2023.
- [8] Dhivya Elavarasan and Durairaj Vincent P M, "Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications", IEEE ACCESS, vol 8, pp.86886-86901, May.2020.
- [9] Ersin Elbasi , Chamseddine Zaki, Ahmet E. Topcu, Wiem Abdelbaki, Aymen I. Zreikat , Elda Cina, Ahmed Shdefat and Louai Saker, "Crop Prediction Model Using Machine Learning Algorithms," Applied Sciences, vol. 13, no. 16, p. 9288, 2023.

- [10] Gautam Gupta, Rashi Setia, Archana Meena, Bhavnesh Jaint, "Environment Monitoring System for Agricultural Application using IoT and Predicting Crop Yield using Various Data Mining Techniques", in Proceedings of the Fifth International Conference on Communication and Electronics Systems (ICCES 2020) pp.1019-1025, 2020.
- [11] Jain S and Ramesh D, "Machine Learning convergence for weather-based crop selection," in 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), IEEE, pp. 1–6, 2020.
- [12] Janmejey Pant, R.P. Pant, Manoj Kumar Singh, Devesh Pratap Singh, Himanshu Pant, "Analysis of agricultural crop yield prediction using statistical techniques of machine learning", in Materials Today: Proceedings, Vol 46, Part 20, pp.10922-10926, Jan.2021.
- [13] Jayanarayana Reddy D, Dr Rudra Kumar M, "Crop Yield Prediction using Machine Learning Algorithm", in Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021), pp.1466-1470, 2021.
- [14] Jeevan Nagendra Kumar Y, Spandana V, Vaishnavi V S, Neha K, Devi V G R R, "Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector", in Proceedings of the Fifth International Conference on Communication and Electronics Systems (ICCES 2020), pp.736-741, 2020.
- [15] Kalimuthu M, Vaishnavi P, Kishore M, "Crop Prediction using Machine Learning", in Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT 2020), pp.926-932, 2020.
- [16] Manjula E, Djodiltachoumy S, "A Model for Prediction of Crop Yield", in International Journal of Computational Intelligence and Informatics, Vol. 6: No. 4, pp.298-305, March.2017.
- [17] Nikhil R, Prof. Anisha B S, Dr. Ramakanth Kumar P, "Real-Time Monitoring of Agricultural Land with Crop Prediction and Animal Intrusion Prevention using Internet of Things and Machine Learning at Edge", in International Conference on Electronics, Computing and Communication Technologies (CONECCT), IEEE, 2020.
- [18] Pachade R S and Sharma A, "Machine learning for weather-specific crop recommendation," Int J Health Sci (Qassim), pp. 4527–4537, Oct. 2022.

- [19] Paul M, Vishwakarma S K, and Verma A, "Analysis of Soil Behavior and Prediction of Crop Yield Using Data Mining Approach," in International Conference on Computational Intelligence and Communication Networks (CICN), IEEE, pp. 766–771, 2015.
- [20] Potnuru Sai Nishant, Pinapa Sai Venkat, Bollu Lakshmi Avinash, Jabber B, "Crop Yield Prediction based on Indian Agriculture using Machine Learning", International Conference for Emerging Technology (INCET), 2020.
- [21] Rakesh Kumar, Singh M P, Prabhat Kumar and Singh J P, "Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique", in International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), pp.138-145, 2015.
- [22] Ranjani J, Kalaiselvi V G K, Sheela A, Deepika Sree D, Janaki G, "Crop Yield Prediction Using Machine Learning Algorithm," in 2021 4th International Conference on Computing and Communications Technologies (IC CCT), IEEE, pp. 611–616, 2021.
- [23] Ravi R and Baranidharan B, "Crop Yield Prediction using XG Boost Algorithm," International Journal of Recent Technology and Engineering (IJRTE), vol. 8, no. 5, pp. 3516–3520, 2020.
- [24] Reuben Varghese, Smarita Sharma, "Affordable Smart Farming Using IoT and Machine Learning", in Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS 2018), pp.645-650, 2018.
- [25] Sai Pravallika G, Sri Thanvi K, Rupa Ch, Kundana L, Sirisha G, "Proficient Smart Soil based IoT System for Crop Prediction", in Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020), pp.752-757, 2020.
- [26] Shripathi Rao M, Singh A, Subba Reddy N V, and Acharya D U, "Crop prediction using machine learning," J Phys Conf Ser, vol. 2161, no. 1, p. 012033, 2022.
- [27] Suresh A, Ganesh Kumar P, and Ramalatha M, "Prediction of major crop yields of Tamilnadu using K-means and Modified KNN," in 3rd International Conference on Communication and Electronics Systems (ICCES), IEEE, pp. 88–93, 2018.

- [28] Teja M S, Preetham T S, Sujihelen L, Christy, Jancy S, and Selvan M P, "Crop Recommendation and Yield Production using SVM Algorithm," in 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE, pp. 1768–1771, 2022.
- [29] Thomas van Klompenburga, Ayalew Kassahuna, Cagatay Catalb, "Crop yield prediction using machine learning: A systematic literature review", in Computers and Electronics in Agriculture, vol 177, 2020.
- [30] Vamsi Krishna, Reddy T, Harsha S, Ramar K, Hariharan S, and Bhanuprasad A, "Analysis of Crop Yield Prediction using Machine Learning algorithms," in 2022 2nd International Conference on Innovative Sustainable Computational Technologies (CISCT), IEEE, pp. 1–4, 2022.
- [31] Veenadhari S, Dr. Bharat Misra, Singh C D, "Machine learning approach for forecasting crop yield based on climatic parameters", in International Conference on Computer Communication and Informatics (ICCCI -2014), 2014.
- [32] Vijayabaskar P S, Sreemathi R, Keertanaa E, "Crop Prediction Using Predictive Analytics", in International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC), pp.370-373, 2017.

## APPENDIX

A review paper titled “Development of a Portable Crop Prediction System with Machine Learning Algorithm for Agriculture” was presented in “1st international conference on Advancement in Energy (URJA Sangam-2023)”, organized by Department of Mechanical Engineering, MNIT Allahabad, India from 18<sup>th</sup> to 20<sup>th</sup> December. The paper has been accepted and is under process for publication.

**Authors:** Dr. Deepak J R, Riswanth S, Devaragothaman Y A.

**Link:** <http://mnnit.ac.in/urjasangam-2023/>

Certificate of Presentation from the conference has been attached below.





The first page of the review paper is attached below.

---

1<sup>st</sup> International Conference on Advancement in Energy (ऊर्जाSangam-2023)  
18<sup>th</sup>–20<sup>th</sup> December, 2023  
MNNIT Allahabad, Prayagraj, Uttar Pradesh, India

## **Development of a Portable Crop Prediction System with Machine Learning Algorithm for Agriculture**

**J R Deepak<sup>a,\*</sup>, Riswanth S<sup>a</sup>, Devaragothaman Y A<sup>a</sup>**

*School of Mechanical Engineering,  
Sathyabama Institute of Science and Technology, Chennai,  
Tamil Nadu, India*

\*E-mail: deepak.prod@sathyabama.ac.in

### **Abstract**

Being in the top three cultivating countries, India is said to be the land of agricultural wealth by nature. Conventional methods of farming have been cherished till now through generations and these practices give life to agriculture. But recent days, environmental factors such as sudden change in climatic conditions, rise in temperature, depiction of soil fertility and humidity affect agricultural growth and thereby decrease the production of crops. Recent machine learning concept, Internet of Things plays a vital role in crop growth. Predicting the crop yield using the historical data and comparing it with the real time data using IOT devices help to increase the productivity of crops more effectively. Based on the recent research and developments in the field of agriculture, an innovative portable crop prediction system for agriculture is developed which collects real time data and predicts environmental factors such as atmospheric temperature, humidity, amount of rainfall, pH level, soil moisture and fertility of the soil. The system consists of a battery, Arduino Mega microcontroller and an array of sensors which include DS18B20 (Temperature Sensor), DHT11 (Humidity Sensor), pH Sensor, NPK Sensor (Soil Fertility Sensor), FC-37 (Rain Sensor), Capacitive Soil Moisture Sensor and Zigbee TX module as its input devices. These sensors contain probes which are inserted into the soil to collect the inputs and fed into the microcontroller. Using Zigbee transmitter and receiver modules, the data is transferred to the IOT system. The data is further transferred to the machine learning algorithm and the results are instantly relayed to dedicated website through Node MCU after the prediction, providing easy access to the users. By integrating the historical and real time sensor data's, the crop prediction system creates a comprehensive environmental profile of the agricultural land using machine learning. This portable crop prediction system provides better sustainability and accurate predictions to the farmers there by increasing their productivity.

Keywords: Crop prediction, Machine Learning, Agriculture, IOT system.

### **1. INTRODUCTION**

Agriculture plays a vital role in ensuring food security, a sustainable economy, and support for livelihoods worldwide. In India, agriculture is not only a key economic sector but also the primary occupation for many people. Table 1.1 illustrates the significant challenges influencing the growth and productivity of crops. The challenges such as meteorological factors, soil factors, biotic factors and hazardous factors make agriculture very difficult for the farmers and demand them to develop innovative solutions to optimize and predict crop yields.

An experimental paper is written, has been accepted for publication and is under process for publication. The first page of the experimental paper has been attached below

## **Portable crop prediction system for crop recommendation based on soil and weather characteristics using machine learning approach for Gen Z farmers**

**J R Deepak<sup>a\*</sup>, Riswanth S<sup>a</sup>, Devaragothaman Y A<sup>a</sup>, Deborah Raj<sup>b</sup>**

<sup>a</sup>*School of Mechanical Engineering, Sathyabama Institute of Science and Technology, Chennai, India*

<sup>b</sup>*Department of Communication, Madras Christian College, Chennai, India*

\*E-mail: [deepak.prod@sathyabama.ac.in](mailto:deepak.prod@sathyabama.ac.in)

[rishidwa@gmail.com](mailto:rishidwa@gmail.com)

[devhimayan@gmail.com](mailto:devhimayan@gmail.com)

### **ABSTRACT**

India is one of the most suitable countries for agriculture, yielding numerous varieties of crops annually. However, in recent days, the condition for agriculture has changed drastically due to the influence of various environmental factors like climatic conditions, nature of soil, pollution levels, natural hazards etc. Addressing these challenges in Indian agriculture, a portable crop prediction system equipped with various sensors is developed for real-time data collection on temperature, humidity, pH, soil fertility, soil moisture, and rainfall. This system leverages machine learning and IOT technology, integrating historical and real-time sensor data collected to accurately predict crucial environmental factors. By implementing MPL classifier, XG Boosting classifier and bagging classifier machine learning algorithms the datasets are trained and tested for its accuracy, thereby finding the most suitable machine learning algorithm for crop prediction. XG Boosting classifier is the most suitable machine learning algorithm for crop prediction. Driven by Arduino Mega and Zigbee modules, the system utilizes the XG Boost classifier algorithm to predict crop yield and recommend suitable crops. The system aims to enhance agricultural productivity by providing farmers with live sensor data display and crop recommendations through a web-based platform, facilitated by Node MCU. This Portable crop prediction system provides farmers with actionable insights, enabling them to adapt to changing environmental conditions and optimize crop productivity. It also offers farmers a comprehensive environmental profile of their agricultural land, enabling informed decision-making and optimization of agricultural practices. This promotes sustainable farming practices and efficiency, resonating with the environmentally conscious mindset of Gen Z farmers.

**Keywords:** Crop Prediction, Machine Learning, IOT, Agriculture, Gen Z farmers

### **INTRODUCTION**

This portable crop prediction system was selected as one of the top innovative projects in the mechatronics department of Sathyabama Institute of Science and Technology, Chennai. As a result, the portable crop prediction system was presented in the project contest and exhibition “Technovation” on 23.03.2024 and was evaluated by industry experts. The poster displaying the portable crop prediction system with the highlights that has been presented in the contest is attached below.



# SATHYABAMA

## INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited with Grade “A” by NAAC | 12B Status by UGC | Approved by AICTE






## PORTABLE CROP PREDICTION SYSTEM WITH MACHINE LEARNING ALGORITHM FOR AGRICULTURE

Department of Mechatronics

**Team Members**

Devaragothaman.Y.A

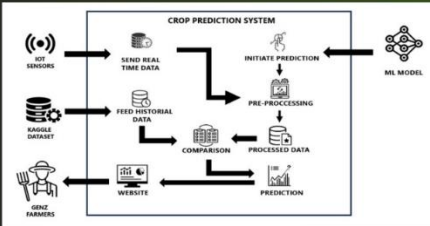
Riswanth.S

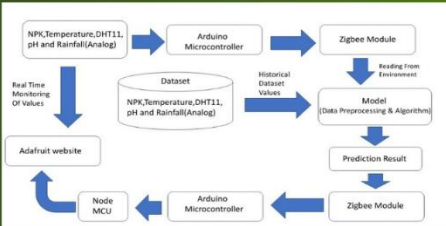
**Project Guide**

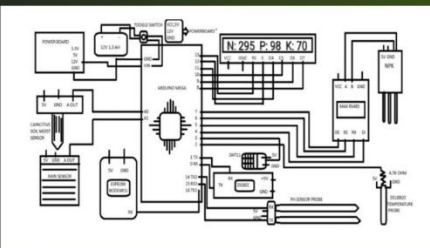
Dr.J.R.Deepak

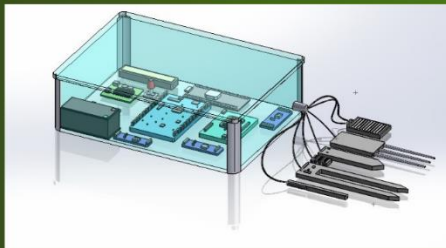
### HIGHLIGHTS OF THE SYSTEM

- The developed portable crop prediction (PCP) device integrates sensors for real-time data collection on various environmental factors like temperature humidity, pH, soil fertility, moisture, and rainfall.
- Sensors collect real-time data, which is processed using machine learning (XG Boost) algorithms which is trained using a historical and real time dataset to predict the crop yield and to recommend the suitable crops.
- PCP system provides live sensor data being display and suitable crop recommendations via a web-based platform.
- PCP enhances decision-making by providing comprehensive environmental profiles of agricultural land, enabling informed choices and optimization of farming practices.
- PCP increases crop yield and quality through precise environmental monitoring and data-driven decision-making.
- The PCP system provides real-time soil nutrient (N, P, K) data, allowing Gen Z farmers to compare recommended nutrient values for predicted crops. This facilitates adjustment of soil nutrients, ensuring optimal conditions for maximum crop yield. Moreover, it also reduces resource wastage by optimizing fertilizer and water usage based on real-time soil and weather conditions.
- This promotes sustainable farming practices and efficiency, resonating with the environmentally conscious mindset of Gen Z farmers.









The plagiarism report of the final report has been attached below

| RE-2022-225282-plag-report |                                                                                                                                                                                                                                                                            |              |                |
|----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|----------------|
| ORIGINALITY REPORT         |                                                                                                                                                                                                                                                                            |              |                |
| 20%                        | 15%                                                                                                                                                                                                                                                                        | 14%          | 9%             |
| SIMILARITY INDEX           | INTERNET SOURCES                                                                                                                                                                                                                                                           | PUBLICATIONS | STUDENT PAPERS |
| PRIMARY SOURCES            |                                                                                                                                                                                                                                                                            |              |                |
| 1                          | sist.sathyabama.ac.in<br>Internet Source                                                                                                                                                                                                                                   | 2%           |                |
| 2                          | fastercapital.com<br>Internet Source                                                                                                                                                                                                                                       | 1%           |                |
| 3                          | www.coursehero.com<br>Internet Source                                                                                                                                                                                                                                      | 1%           |                |
| 4                          | Submitted to University of Sunderland<br>Student Paper                                                                                                                                                                                                                     | 1%           |                |
| 5                          | N. V Suresh Krishna, Pokala Neeraj, Bathina Surya, B.V. Prudhvi, V. Hima Deepthi.<br>"Machine Learning Algorithms for Crop Yield Prediction in Real-Time Scenarios", 2023 4th International Conference on Signal Processing and Communication (ICSPC), 2023<br>Publication | <1%          |                |
| 6                          | Tetala, Satya Surya Dattatreya Reddy.<br>"Artificial Intelligence Powered Personalized Agriculture", Colorado State University, 2023<br>Publication                                                                                                                        | <1%          |                |
| 7                          | hackaday.io                                                                                                                                                                                                                                                                |              |                |

# RE-2022-225282 - Turnitin Plagiarism Report

*by Riswanth S*

---

**Submission date:** 30-Mar-2024 01:16AM (UTC+0300)

**Submission ID:** 271711770685

**File name:** RE-2022-225282.docx (8.54M)

**Word count:** 12984

**Character count:** 73238