IOT BASED SOIL FERTILITY ANALYSIS WITH ML

***ABSTRACT:***

***Farmers need to predict soil nutrient levels to optimize crop productivity and ensure a surplus of high-quality crops, especially with the increasing demand for food and growing population. This requires advanced technologies to assist farmers in monitoring their crops and making informed decisions about crop management practices.To address this need, smart agro devices have been developed with IoT sensors that can measure key soil criterion such as moisture content, torridity, pH levels, and nutritional levels. These devices send soil-related information to farmers and agriculture experts on an hourly basis and transmit the collected data wirelessly to a central database or cloud platform for analysis using ML algorithms.Using ML algorithms to analyze the data, farmers and agriculture experts can obtain insights into the current state of the soil and predict future nutrient levels. Also the data here can then be used to make a conclusion about crop management practices such as fertilization, irrigation, and crop rotation. With the timely monitoring provided by smart agro devices, farmers can take preventive measures against soil nutrient deficiencies or excesses, leading to increased crop yields and better crop quality.In this study, smart agro devices with IoT sensors and ML algorithms offer a powerful tool for improving crop yields and producing healthy crops. By providing timely and accurate information on soil conditions, these devices can help farmers optimize crop management practices, contributing to the nation’s economy and progress.***

***Keywords: soil, crop, moisture, humidity, water, IOT sensors, Machine learning, decision tree, naive bayes.***

1. INTRODUCTION

The cornerstone of efficient agricultural practices lies in the vitality of the soil. The presence of vital nutrients in a soil that are necessary for the development of plants makes it a fertile soil. Additionally, physical characteristics of soil structure and aggregation play a vital role in enabling roots to explore and allowing air and water to penetrate. The terms "soil health" and "soil quality" can be used interchangeably when describing soils that are not only rich in nutrients but also possess beneficial physical and biological properties. The quality of soil is measured by how effectively it meets our requirements.Soil quality can be defined as the capacity of a particular soil type to facilitate the growth and Fostering the growth of flora and fauna, enhancing the purity of water and air, and offering a foundation for human habitation and well-being are all benefits that can be derived from maintaining or improving soil productivity.To cultivate robust and high-quality crops, it is essential that the soil contains 45% nutritional deposits, 5% natural substances, 25% oxygen, and 25% liquid particles. The nutrient content of the soil is a critical factor in determining crop selection, growth, and overall productivity. Saying that the agricultural industry is in difficulty and cries out with the rest of the economy seems insufficient. Crop yield and productivity need to be improved. Agriculture is the science and practise of raising plants, animals, and other living things. It has existed for aeons, long before early human civilization was founded by humans. Plants require an abundance of water, air, and sunlight. The needed information will be obtained by measuring the various factors. Measuring the soil’s electrical conductivity is one technique to keep track of all these important nutrients. Studies on electrical conductivity reveal the need for the proper dietary requirements. Measuring ions and their temperature-dependent activity is vital in ascertaining the electrical conductivity of soil. Additionally, the porosity and cation exchange capacity (which are linked to soil texture) that affect electrical conductivity will also be gauged.Electrical conductivity and pH testing provide information about the nutrients that are available to plants and their real needs.Soil fertility refers to the soil’s ability to provide the necessary nutrients for crop growth. The primary nutrients that plants obtain from the soil are nitrogen, phosphorus, In order to promote the healthy growth of crops, it may be necessary to supplement the soil with fertilizers, manure, or compost to provide essential nutrients such as potassium, calcium, and magnesium. Although there are many other nutrients that plants acquire from the soil, typically the secondary nutrients are already present in sufficient quantities, so additional supplementation is unnecessary. The pH level of the soil is also a critical aspect of soil fertility, as it affects plant growth. Soil pH measures soil acidity and does not provide nutrients to plants. Most crops thrive when the soil pH is between 6.2 and 6.8.

1. RELATED WORKS

Machine learning and sensing technologies are being used for crop yield prediction and nitrogen status estimation in precision agriculture. The analysis of soil reaction is crucial. Further advancements are expected in the future, such as targeted sensor application, combining different sensor methods and technical expertise, and the development of hybrid systems combining machine learning and signal-processing methods. These developments have the potential to transform crop production monitoring[1]. IoT can reduce food loss and agricultural inputs in agriculture, but large-scale pilots are needed for adoption. Challenges include interoperability, business models, privacy, safety, and data governance. The paper outlines scientific and agricultural criteria to address these issues and presents an architecture utilizing a system approach. The pilot covers dairy, fruit, crops, flesh, and plants. A cultural shift is required for successful implementation of IoT in agriculture[2]. Machine learning models are being integrated with agrarian factors for yield forecasting, focusing on soil temperature and electrical conductivity parameters. These models help to maximize crop yield by fusing multiple data resources and analyzing them using various analytical models like Decision Trees, random forest models, Support Vector Machines, Bayesian Networks, and artificial neural networks. They facilitate precision farming and water management by analyzing soil and climate data[3]. Data mining techniques are used for soil and plant nutrient management in organic crops, focusing on the analysis of soil nutrients and their impact on crop cultivation. Soil quality is crucial for determining crop productivity, and data mining tools are used to analyze the nutrients in soil, particularly in a few areas of Tamil Nadu, India. Different types of soil contain varying concentrations of nutrients such as phosphorus, potassium, magnesium, calcium, nitrogen, sulphur, iron, and zinc. The study utilizes a mixed neural network method to investigate soil nutrient levels[4]. Comparison of machine learning models for predicting monthly soil temperature at different depths found ELM to be the most accurate. Soil temperature impacts agricultural processes and soil health, and varies with depth. Air temperature measurements suffice for depths of 5, 10, and 50 cm, but for 100 cm, radiation and wind speed data are necessary. Soil temperature affects plant development, root environments, and soil-based microorganisms. Measuring and analyzing soil temperatures at different depths is crucial for plant growth and soil health[5]. Study used Extreme Learning Machine (ELM) to categorize and predict village-level soil parameters for optimizing crop growth and reducing fertilizer input costs. Gaussian radial basis function was found to perform the best for four out of five issues, while hyperbolic tangent algorithm was most accurate for pH classification. This approach has the potential to improve soil health, increase profitability, and enhance environmental quality[6]. Data mining techniques were used to predict crop yields and explore their application in agriculture. Uncertain sets were utilized for approximative answers to problems with missing data patterns and information interactions. Various data mining techniques were examined, including GPS, to identify important information related to agriculture, such as soil identification[7]. Model Predictive Control (MPC) techniques are being implemented in agriculture to increase productivity and efficiency in irrigation systems, agricultural machinery, production, and greenhouses. MPC can handle complex nonlinear systems with huge time delays and has shown promising results. However, the latest MPC techniques have not been widely used in agriculture, and further study and improvement are necessary to fully evaluate its advantages[8]. A hybrid approach combining rough set on fuzzy approximation space and neural network is developed to predict crop suitability based on natural resources. The model achieves a validation classification accuracy of 93% when tested on agricultural data from Tamil Nadu's Vellore Region. The method converts quantitative data into qualitative data using rough set and predicts the choice attribute value using a back propagation neural network. The approach aids farmers in decision-making regarding crop types to cultivate on their farms[9]. Automation using IoT sensors and AI can increase agricultural productivity by providing real-time data on crop development. This study proposes a smart farming technology using sensors for temperature, humidity, moisture, pH, and level. The Multi-Layer Perceptron (MLP) method is used to determine the crop that can be produced in the soil. The IoT hardware comprises a microcontroller linked to diverse sensors, generating a real-time dataset for validation, which is then subjected to the MLP Neural Network algorithm[10].

1. SYSTEM STUDY

OVERVIEW

IoT (Internet of Things) based soil fertility analysis with Machine Learning (ML) involves the use of various sensors, IoT devices, and ML algorithms to monitor and analyze soil health parameters. This technology is used to optimize agricultural productivity by providing precise and real-time information about soil fertility, crop health, and environmental factors. The IoT-based system has a number of sensors, including ones for moisture, temperature, humidity, and nutrients, which gather information from the soil. IoT devices like gateways, routers, and other network devices are used to transmit the acquired data to a central database. After the data has been gathered, machine learning algorithms are used to interpret it, providing real-time analysis of soil health factors including nutrient content, and soil moisture content. The acquired data is used by the machine learning algorithms to build prediction models that can pinpoint the ideal planting techniques, crop rotation schemes, and soil management procedures. Additionally, with the aid of this technology, farmers can choose wisely when to apply fertilisers and other soil additives, improving yields and minimising environmental effect.Reduced expenses, increased yields, and less environmental effect are all advantages of IoT-based soil fertility analysis using ML. Additionally, with the use of this technology, farmers may see issues with their crops before they worsen, enabling them to take preventative measures to reduce future crop losses. Additionally, it may assist farmers in streamlining their irrigation schedules, which promotes water conservation and lower water use. IoT-based fertilisation assessment with ML, in conclusion, is an exciting innovation that has the potential to completely change how we manage our crops. It may assist farmers in making knowledgeable choices about their crops, enhancing production and minimising environmental impact.

OBJECTIVE

In order to forecast soil characteristics including macro- and micronutrients, soil reactions, carbon from organic materials, and electrical conductivity, this article examines the use of several multivariate statistical data analysis approaches and machine learning algorithms. These potent instruments allow for precise forecasting and new understanding of critical soil features, both of which are important for maximising agricultural yields and preserving soil quality. Obtaining and creating plans for sensors to evaluate soil characteristics The goal of this project is to develop a homegrown Internet of Things (IoT) gadget capable of measuring and validating a wide range of soil physical properties, from micro- to macro-nutrients.

PROBLEM STATEMENT

The acquired data is used by the machine learning algorithms to build prediction models that can pinpoint the ideal planting techniques, crop rotation schemes, and soil management procedures. Additionally, with the aid of this technology, farmers can choose wisely when to apply fertilisers and other soil additives, improving yields and minimising environmental effect.Reduced expenses, increased yields, and less environmental effect are all advantages of IoT-based soil fertility analysis using ML. Additionally, with the use of this technology, farmers may see issues with their crops before they worsen, enabling them to take preventative measures to reduce future crop losses. Additionally, it may assist farmers in streamlining their irrigation schedules, which promotes water conservation and lower water use. IoT-based fertilisation assessment with ML, in conclusion, is an exciting innovation that has the potential to completely change how we manage our crops. It may assist farmers in making knowledgeable choices about their crops, enhancing production and minimising environmental impact.

PROPOSED SOLUTION

Utilising multivariate data analysis methods, machine learning classifiers, and the Internet of Things, this research seeks to foretell the supply of micro and macronutrients, soil reactivity, and other critical physical parameters in agriculture. This will help farmers understand more about the crops they grow and make any necessary improvements. The collected data will be used to train machine learning classifiers to make predictions about soil reactivity, micro- and macronutrient availability, and other physical parameters. These representations will be created using both current and past information gathered by Internet of Things gadgets. To ensure the reliability of the prediction models, the research will entail validating the soil’s macro- and micronutrients, reaction, conductivity to electricity, and other physical qualities using traditional data from agricultural departments. The researchers may then evaluate how well their predictions matched reality and adjust their models accordingly. As a whole, the project hopes to advance sustainable agriculture by equipping farmers with multivariate data analysis methods, Machine Learning classifiers, and the Internet of Things to improve the quality and yield of their crop harvests.

1. PROPOSED SYSTEM

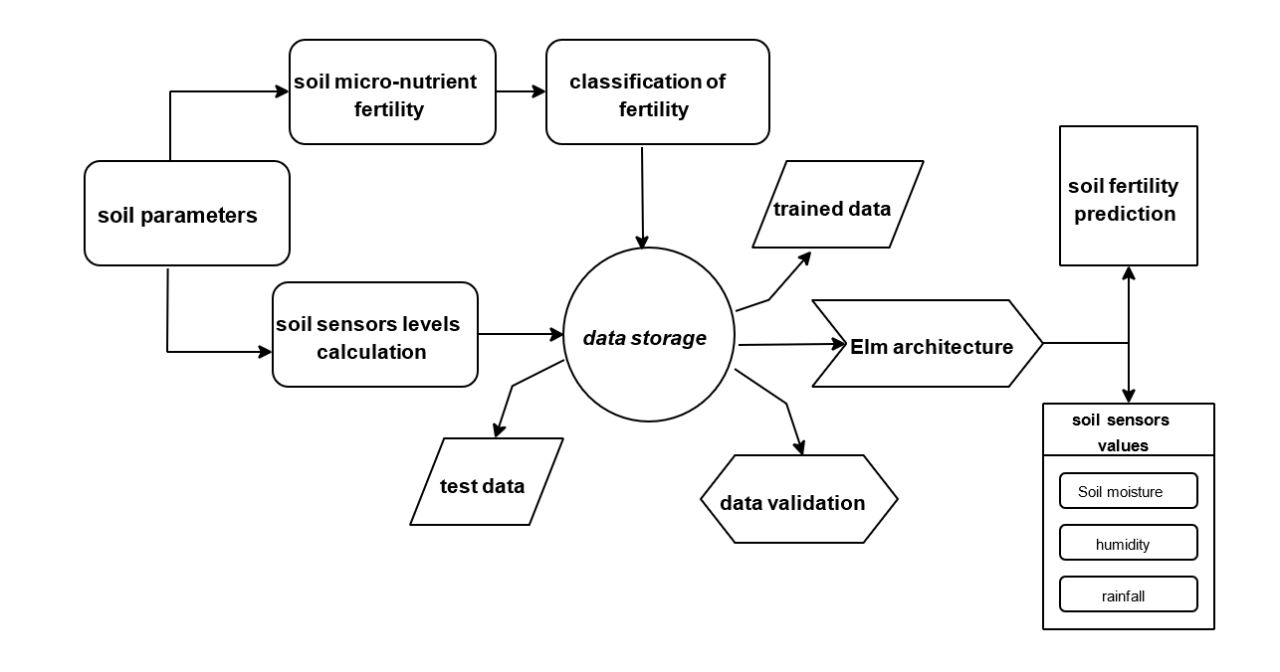


Fig.1 ARCHITECTURE DIAGRAM

Smart Agro Device

The smart agro device is essentially a sensor node that can be embedded into the soil to monitor various soil parameters such as pH level, moisture content, temperature, and nutrient levels. The device should be equipped with sensors that are calibrated to measure these parameters accurately.The device can be powered by a battery, solar power, or an external power source. The data collected by the device is then transmitted to a central database for further analysis.

Data Acquisition

The data acquired by the smart agro device is preprocessed in real-time to ensure that it is accurate and consistent.The preprocessing may include cleaning the data to remove any outliers or errors, filtering the data to eliminate noise or interference, and normalizing the data to ensure that it is on the same scale. The preprocessed data is then stored in a database for further analysis.

Machine Learning Model

A ml model can be utilized to predict soil nutrient levels based on historical data that has undergone preprocessing. The selection of a specific ml algorithm, such as decision trees, random forests, or support vector machines, is determined by the nature of the data and the precision demands of the system. The training of the model involves utilizing a substantial dataset of historical soil information that has been annotated with corresponding nutrient levels.

Predictive Analytics

The predictive analytics component uses the trained machine learning model to predict the soil nutrient levels in real-time. The predicted values are then sent back to the smart agro device for display to farmers and agriculture experts.The predictive analytics component may also include a data visualization tool that enables users to view the nutrient levels graphically over time.

User Interface

The user interface component provides a user-friendly interface for farmers and agriculture experts to view the real-time soil nutrient levels and receive alerts in case of any abnormalities. The user interface can be a mobile application, a web application, or both. The user interface should provide real-time updates of the soil nutrient levels, as well as historical trends and other relevant information.

Cloud Infrastructure

The cloud infrastructure provides a scalable and secure platform for hosting the IoT-based soil fertility analysis system. The cloud infrastructure may include a cloud database for storing the preprocessed data, a machine learning platform for training and deploying the machine learning model, and a cloud application platform for hosting the user interface. The cloud infrastructure should be able to handle a large volume of data and provide high availability and reliability. The IoT-based soil fertility analysis system with machine learning is a sophisticated system that combines sensor technology, machine learning, and cloud computing to enable real-time monitoring and analysis of soil nutrient levels. The system has the potential to revolutionize agriculture by improving crop productivity and quality while reducing the environmental impact of farming.

Elm Architecture

The Elm Architecture is a software design paradigm used to build interactive applications like online tools and video games. What’s unique about this approach is that it seems to develop organically, rather than being intentionally designed. As more and more programmers started using Elm, they noticed that they were using the same fundamental patterns repeatedly in their code. These patterns became the foundation for what we now call the Elm Architecture. So instead of someone coming up with this architecture and imposing it on developers, it emerged naturally as a best practice through the collective experience of the Elm community.

Training data

Training data refers to the set of data used to train an algorithm or machine learning model to predict outcomes. If you’re using supervised learning or a hybrid approach that includes this method, your data will typically be enhanced with data labeling or annotation. This is done to provide additional information to the model and guide its learning process.

Testing data

Once you have built your machine learning model using the training data, you need to evaluate its performance by testing it with unknown data. This testing data allows you to assess the effectiveness of your algorithms and identify areas for improvement or optimization to achieve better outcomes. Therefore, testing data is crucial in the machine learning process as it helps to validate the accuracy and robustness of your model.

1. METHODOLOGY

Smart Agro device using IoT

a) ESP 8266:The ESP8266 is a microcontroller chip that comes with a complete TCP/IP stack and the ability to connect to Wi-Fi networks. It is programmable using the Arduino IDE and has the capacity to operate independently or be linked to a microcontroller. 6.1.2

Espressif

A low-cost microprocessor known as ESP8266 is manufactured by Systems, which is based in Shanghai, China. It is designed with integrated TCP/IP networking software, microcontroller capabilities, and the ability to connect to Wi-Fi networks 6.1.3

The ESP-01 module

During August 2014, the ESP8266 chip, which was manufactured by Ai-Thinker, an independent maker, gained traction in the English-speaking maker community. This chip enable microcontrollers to easily connect to Wi-Fi networks and establish uncomplicated TCP/IP connections by using Hayes-style commands. However, there was initially a dearth of information available in English about the chip and its functionalities. 6.1.4



Fig.2 ESP MODULE

Processor

The Tenisilica Diamond Standard 106 Micro is a 32bit RISC CPU core that operates at 80 or 160 MHz. It is based on the L106 microarchitecture and is often used in microcontroller applications, including the ESP8266 Wi-Fi microprocessor. The Tensilica Diamond Standard 106Micro is known for its energy efficiency and can handle various types of data processing, including audio, video, and image processing. 6.1.5

Memory

the ESP8266 has 32 kilobytes of Instruction RAM available for storing the program code that it needs to execute. This Instruction RAM can be accessed quickly by the ESP8266’s CPU, allowing it to execute code with minimal delay

Instruction

The chip has a cache memory of 32 KiB RAM, with an additional 16 KiB ETS system-data RAM and 80 KiB user-data RAM. The ESP8266 also supports outer Q.S.P.I flash memory of up to sixteen MiB. (five hundred twelve KiB to four MiB typically included)

Wi-Fi

The ESP8266 includes a Wi-Fi transceiver that is compatible with IEEE 802.11 b/g/n standards. It has a built-in T.R switch, Balunn, lna, and power amplifying device, and supports authentication for W.E.P or W.P.A/W.P.A2 networks, as well as open networks. The chip comes equipped with 17 GPIO pins that support both SPI and I2C serial peripheral interface buses through software implementation. I2S interfaces are also supported, along The ESP8266EX includes a GPIO2 pin that can be used to enable a transmit-only UART. In addition, the chip features a 10-bit ADC and a DMA UART on dedicated pins, with a successive approximation ADC.



Fig.3 WiFi module

High Durability

The ESP8266EX is well-suited for use in industrial settings thanks to its ability to operate across a wide temperature range. It is a reliable, compact, and sturdy chip that comes with fully integrated on-chip features and requires only a small number of external discrete components. 6.1.9

Compactness

The ESP8266EX is a tiny yet powerful chip that integrates a 32-bit Tenisilica® processor, digital peripheral interfaces, RF baluns, power amplifiers, low noise receive amplifiers, filters, antenna switches, and power management modules. All these on-chip functionalities make the ESP8266EX a highly integrated and robust solution for various IoT applications. 6.1.10

Power-Saving Architecture

The ESP8266EX chip is designed specifically for mobile devices, wearable electronics, and Internet of Things (IoT) applications, and incorporates several technologies to minimize power consumption. It supports three modes of operation, including active mode, sleep mode, and deep sleep mode, all of which are integrated into its power-saving architecture, enabling battery powered devices to operate for extended periods. 6.1.11

32-bit Tensilica Processor

The ESP8266EX microcontroller features the Tenisilica L106 32-bit RISC processor, optimized for ultra-low power consumption and capable of achieving a maximum speed of 160 MHz. The RTOS and Wi-Fi stack offer developers access to roughly 80% of the processing power to create user application programs.

Soil moisture sensor

The soil moisture sensor is an inexpensive electrical device that measures the moisture content of soil. It can provide an estimate of the soil’s volumetric water content. The sensor consists of two primary components: the sensing probes and the sensor module. The soil moisture sensor consists of two main components: the sensing probes and the sensor module. When inserted into the soil, the probes allow a current to pass through and measure the resistance value, which corresponds to the moisture content of the soil. The data collected from the probes is then transmitted to the sensor module for processing which then processes it and outputs the analysis as a digital or analogue signal. Consequently, the Soil Moisture Sensor may output both digitally (DO) and analogly (AO). (AO).



Fig. 4 Soil Temperature Sensor

The sensor that measures soil temperature and humidity is the cornerstone of agricultural operations and is linked to crop development. The soil temperature sensor can monitor the temperatures of the soil, atmosphere, and water for experiments and scientific research. 6.1.14

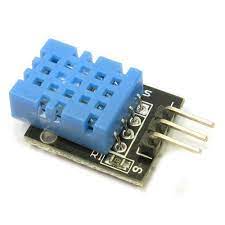


Fig.5 SOIL TEMPERATURE SENSOR

LCD Display

To operate an LCD display using a microcontroller, multiple interface pins must be manipulated simultaneously due to the parallel interface used by LCDs. In the LCD interface, there is a pin called register select (RS), which plays a crucial role in specifying the position in the LCD’s memory where data is to be written. The LCD’s controller has the flexibility to search for instructions on the next operation from the instruction register or save the data to be displayed on the screen in the data register. The choice between these options is left to the user to decide. The Read/Write (R/W) pin serves the purpose of selecting either the read or write mode of the LCD, while the Enable pin enables writing to the 8 data pins of the registers (D0-D7). The high and low states of these pins correspond to the bits or values being written to the registers. During the reading process, the same pins are used to retrieve data from the LCD. Furthermore, the power supply pins (+5V and GND), LED backlight pins (Bklt+ and Bklt-), and display contrast pin (Vo) can be utilized to supply power to the LCD, regulate the display contrast, and activate or deactivate the LED backlight. As a part of the display control process, the data required for the desired image are stored in the data registers, while instructions are loaded into the instruction register. The LiquidCrystal Library simplifies this process by providing high-level commands, eliminating the need for knowledge of the low-level commands. Hitachi-compatible LCDs offer two control modes: 4-bit and 8-bit LCD displays that use the Hitachi HD44780 driver can be easily controlled using the LiquidCrystal library, regardless of the specific model. These displays typically have a 16-pin interface. To control an LCD display using a parallel interface, the microcontroller needs to manipulate multiple interface pins concurrently. The interface of an LCD display comprises several pins that need to be concurrently controlled by the microcontroller to operate the display in parallel. These include the register select (RS) pin, which determines the memory location in the LCD where data is written. The user can choose between using the instruction register for the LCD’s controller to search for the next operation to execute, or the data register to store the information that will be displayed on the screen.

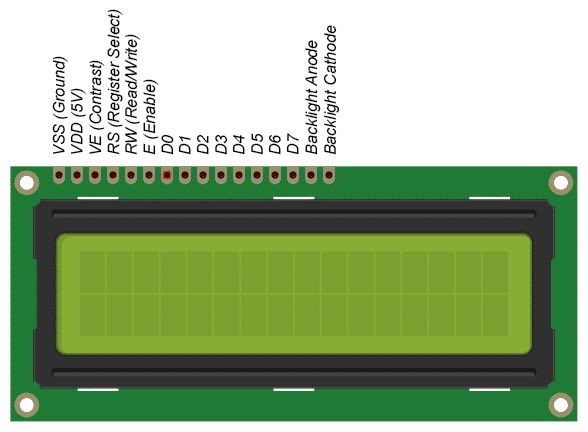


Fig.6 LCD DISPLAY

Machine learning techniques

ml classifiers Machine learning techniques can be leveraged for soil mapping, prediction, and to anticipate soil behavior. These techniques involve training models to analyze various soil properties and make accurate predictions based on that analysis. By using machine learning, we can gain a better understanding of the complex behavior of soil and make more informed decisions about land use, agriculture, and environmental management. 6.2.1

Stochastic Gradient Descent

Stochastic Gradient Descent (SGD) is a popular optimization algorithm in machine learning, which is utilized to determine the optimal parameters for a given model to minimize the loss function. The SGD algorithm works by iteratively adjusting the model parameters in the direction of the negative gradient of the loss function to minimize it. In the SGD algorithm, the term "stochastic" refers to the approach of computing the gradient for a random subset of the training data or even for a single data point, instead of the entire dataset. This method enables the algorithm to be computationally efficient and capable of handling large datasets

Naive Bayes algorithm

Naive Bayes is a popular supervised learning algorithm used for classification problems that is based on the Bayes theorem. Although it can be applied to other data types as well, Naive Bayes is particularly advantageous for text classification tasks, including sentiment analysis, spam detection, and topic classification. The Bayes theorem is a mathematical formula that allows us to calculate the probability of an event occurring given some prior knowledge or evidence. Naive Bayes assumes that the features of a dataset are independent of each other, which makes the calculations simpler and faster than other methods. Naive Bayes is a probabilistic model, meaning it calculates the probability of a data point belonging to each class and then selects the class with the highest probability as the prediction. It achieves this by first learning the probabilities of each feature for each class from the training data. These probabilities are then used to calculate the probability of a new data point belonging to each class.

Decision Tree

Decision Tree is a popular supervised learning algorithm that can be used for both classification and regression tasks. It is a tree-based model that builds a flowchart-like structure to represent a set of decisions that leads to a final outcome. it’s starts with a single node, called the root node, which represents the entire dataset. From there, the algorithm recursively splits the dataset into smaller subsets based on some condition or criteria, such as the value of a particular feature. Each internal node in the tree represents a decision based on a feature or attribute, and each branch represents a possible outcome of that decision. Finally, the leaf nodes represent the final output or outcome of the decision tree.

Support Vector Machine

While SVM is suitable for both classification and regression tasks, it is a robust supervised learning algorithm mainly used for classification purposes. SVM’s primary aim is to locate the optimal hyperplane that can divide different class data points within the feature space. The hyperplane is chosen such that the margin between it and the nearest data points of each class is maximized. This margin represents the distance between the hyperplane and the closest data points, and SVM seeks to maximize this distance to determine the best decision boundary. To handle non-linearly separable data, SVM employs a technique known as kernel trick, which maps data points to a higher-dimensional space where linear separation is possible. This approach enables SVM to effectively address complex and non-linear data. The dimension of the hyperplane is determined by the number of features present in the data. For instance, if the data has two features, the hyperplane is a line, while if it has three features, the hyperplane is a plane, and so on.

KNN algorithm

The KNN algorithm is a type of supervised learning classifier that predicts the classification of a new data point by considering the proximity of other data points. It is known as a nonparametric and instance-based algorithm because it doesn’t require any assumptions to be made about the underlying data distribution. Instead, it uses the training data to determine the classification of new data points based on their similarity to other data points in the dataset. To predict the class of a new data point, the KNN algorithm first calculates the distances between the new data point and all of the training data points in the dataset. It then selects the K training examples that are closest to the new data point based on the distance metric used. The value of K is a hyperparameter that determines the number of nearest neighbors to take into account when making a prediction. Once the K nearest neighbors are identified, the algorithm assigns the class that is most common among those K neighbors as the predicted class for the new data point. When it comes to classification, after identifying the K nearest neighbors, the KNN algorithm examines their labels and determines the most frequent one among them. This most frequent label is predicted as the label for the new data point. However, in regression problems, the KNN algorithm computes the average of the labels of the K nearest neighbors to predict the value for the new data point.

ThingSpeak

Connecting sensors and other internet-enabled devices to the cloud has never been easier than with ThingSpeak, an open-source IoT platform. It was created in Ruby by ioBridge, which introduced it in 2010 to facilitate IoT applications. ThingSpeak’s provides a simple and easy-to-use API that enables devices to send data to the cloud for storage and analysis. The platform allows users to log and analyze data from various sources, including sensors, social networks, and web services. Users can also share data and collaborate with other users to create custom applications and solutions. ThingSpeak’s real-time data visualization and analysis capabilities are among its key features. The platform provides built-in visualization tools and analytics capabilities that allow users to create custom dashboards and graphs to monitor and analyze their data. Users can also

create alerts and notifications based on specific events or thresholds. ThingSpeak is a free and open-source software, which makes it accessible to a wide range of users and developers. It is also highly customizable, allowing users to modify and extend the platform to meet their specific needs.

Overall, ThingSpeak is a powerful and versatile platform that makes it easy for users to connect and manage their IoT devices and sensors. Its simplicity, flexibility, and community-driven development make it an attractive option for both hobbyists and professionals working in the IoT space.

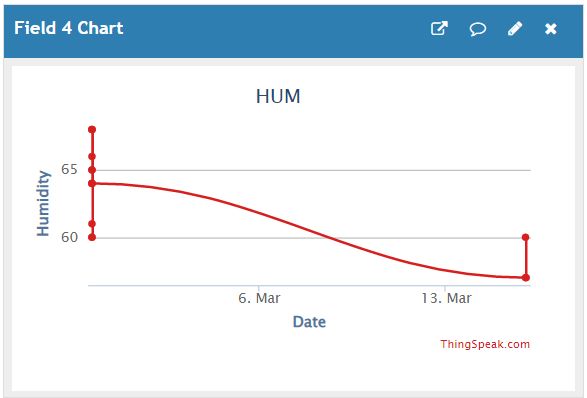


Fig.7 SOIL MOISTURE DATA CHART

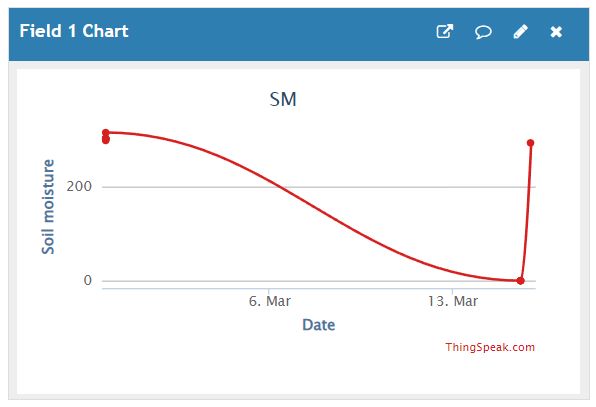


Fig.8 WATER LEVEL DATA CHART

The above graphs are plotted to show the variation in data from Thingspeak. ThingSpeak, a platform for IoT data collection and analysis, has been seamlessly integrated

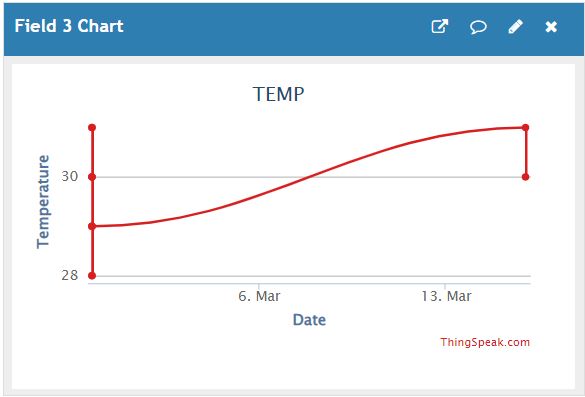


Fig.9 SOIL TEMPERATURE DATA CHART



Fig.10 SOIL HUMIDITY DATA CHART

with the popular mathematical computer program MATLAB by MathWorks. By integrating with MATLAB, ThingSpeak users can analyze and visualize their uploaded data using MAT LAB without requiring a MathWorks MATLAB license. ThingSpeak’s has garnered attention from various "Maker" websites, such as Instructables, Codeproject, and Channel 9, which have featured stories about the platform. ThingSpeak is a platform for IoT analytics that simplifies the collection, visualization, and analysis of real-time data streams. Without the need for any coding, users can send data from their devices to the platform and visualize it instantly. ThingSpeak facilitates the hassle-free monitoring of real-time data streams.

Thingspeak Setup

To get started with ThingSpeak, you can either sign in with your existing account or create a free MathWorks account. Once you’re logged in, select the specific ThingSpeak channel that you want to receive data streams from. If you need assistance in setting up a new channel, refer to the guide on "Collecting Data in a New Channel". For the selected channel, make sure to take note of the following details: In the channel view, the channel ID is displayed as the topmost item, and it’s important to note down the API key located on the API Keys page of the channel view.

Jupyter Notebook

Jupyter Notebook, previously known as IPython Notebook, is a web-based interactive computational environment that facilitates the creation of notebook documents. Its development is supported by various open-source libraries, such as IPython, Tornado, ZeroMQ, jQuery, MathJax, and Bootstrap, which enable the necessary functionality and features for notebook document creation within the Jupyter Notebook environment.

Python Packages Used

While it is possible to perform basic mathematical operations in Python without the use of specialized packages, using the NumPy library can greatly simplify the process of performing complex mathematical operations. The NumPy library provides a wide range of mathematical functions and operations, including matrix operations, linear algebra, Fourier transforms, and more. By using NumPy, users can avoid the need to write complex mathematical algorithms from scratch, saving time and improving code efficiency. Overall, NumPy is an essential library for anyone working with complex data and mathematical operations in Python.

NumPy

provides a range of tools and functions that enable users to easily create and manipulate multi-dimensional arrays, and perform calculations on the data they contain. With NumPy, users can perform a wide range of tasks, such as solving algebraic equations, carrying out statistical analysis, and much more. The library includes a variety of functions for performing basic and advanced mathematical operations, as well as tools for working with complex data structures such as matrices and tensors. Overall, NumPy is a powerful and essential library for any data scientist or developer working with numerical data in Python.

Pandas

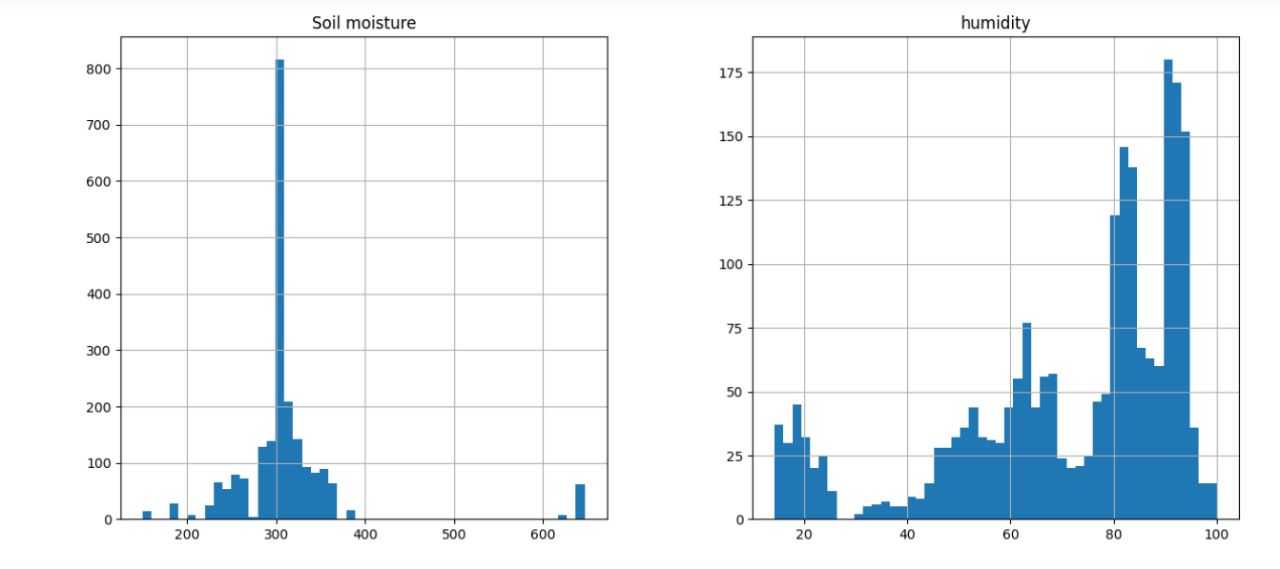
Python’s Pandas library is designed to offer user-friendly data analysis tools and data structures with considerable power.

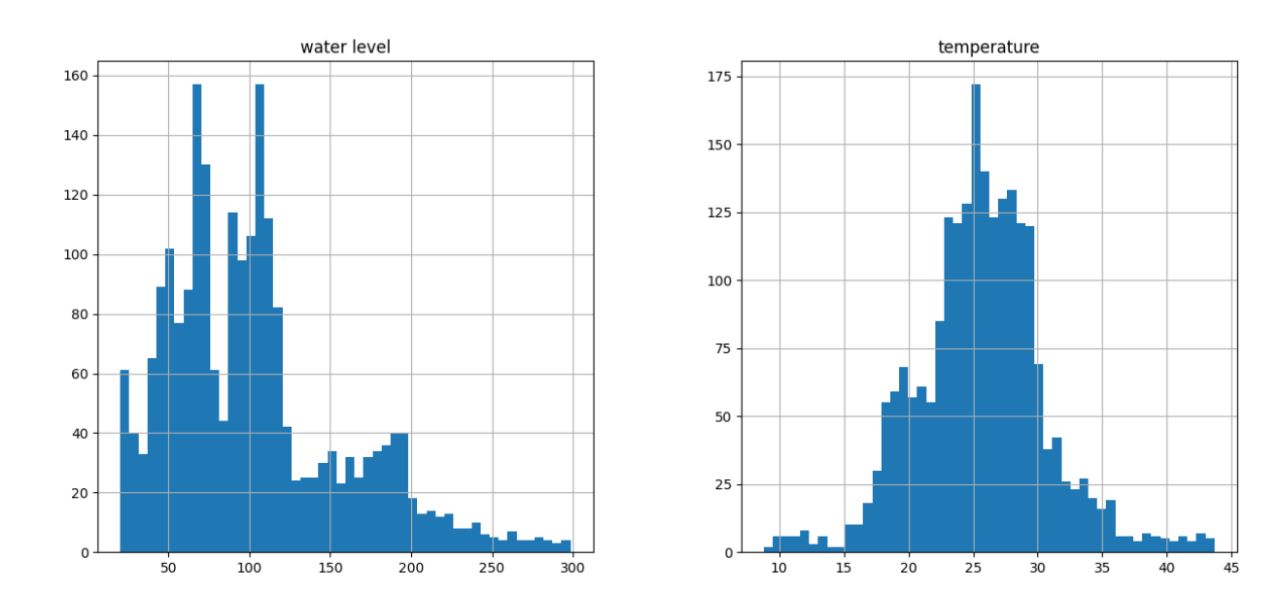
Matplotlib

Matplotlib is a powerful and versatile two-dimensional plotting library that can create interactive, highly efficient visualization charts in a variety of formats. It is cross-platform, meaning that it can be used on various operating systems, including Windows, macOS, and Linux, and can be used with a wide range of programming languages, including Python, C++, and MATLAB.With Matplotlib, you can easily customize plots to meet specific requirements, including adding labels, titles, legends, and annotations.

1. RESULTS AND DISCUSSION

On the basis of the findings that were acquired from the various algorithms, it is possible to draw the conclusion that the algorithms known as Stochastic Gradient Descent (SGD) and Decision Tree are the ones that provide the best results. Both of these algorithms achieved an accuracy score of 90% and an F1 score of 0.9. Additionally, the Naive Bayes algorithm did quite well, with an accuracy score of 89.1% and an F1 score of 0.891. Two common models used in machine learning are Support Vector Machines (SVM) and K-Nearest Neighbor (KNN) did not perform as well, attaining accuracy ratings of 67.3% and 49.5%, respectively. When taking into account the levels of accuracy and F1 scores, the Decision Tree and SGD models emerge as the most promising candidates for this specific dataset. However, it is essential to keep in mind that the optimal model for a given dataset may vary depending on the nature of the data as well as the particular challenge that is being tackled. When deciding which algorithm is most suited to solve a certain issue, in addition to accuracy and F1 scores, it is important to take into account a variety of other aspects, including the length of time required for training and prediction, the ease with which the model can be interpreted, and the overall level of model complexity.





1. CONCLUSION

In this study I have derived my own dataset from soil collected from our surroundings and the highest accuracy in predicting the soil fertility is given by the Decision Tree algorithms often comes out on top. In the context of predicting soil fertility, it makes sense that this algorithm would perform well since it is capable of handling both categorical and continuous data.The interpretability of the decision tree is an advantage.which means that it can provide insights into which features are most important in determining the outcome.The application of machine learning classifiers and multivariate statistical analysis techniques to soil mapping and the prediction of soil behavior has the potential to revolutionize the agricultural industry. With the development of cost-effective and highly efficient smart agro devices, farmers and agricultural industries can better predict soil behavior, which in turn can lead to better yields and improved agricultural production.The use of artificial intelligence techniques in the design of smart agro devices enables the devices to accurately predict soil behavior, taking into account multiple variables and factors that influence soil quality. These devices can provide farmers with realtime information about soil moisture levels, nutrient content, and other important parameters, enabling them to make better decisions about crop management and yield optimization.The implementation of predictive analysis using machine learning classifiers and multivariate statistical analysis techniques can greatly benefit many districts in the country. By accurately predicting soil behavior, farmers can optimize crop yields, reduce crop damage due to pests.

1. REFERENCE
2. Chlingaryan, A., Sukkarieh, S. & Whelan, B. (2018), ‘Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review’, Computers and electronics in agriculture 151, 61–69
3. Brewster, C., Roussaki, I., Kalatzis, N., Doolin, K. & Ellis, K. (2017), ‘Iot in agriculture: Designing a europe-wide large-scale pilot’, IEEE communications magazine 55(9), 26–33.
4. Elavarasan, D., Vincent, D. R., Sharma, V., Zomaya, A. Y. & Srinivasan, K. (2018), ‘Forecasting yield by integrating agrarian factors and machine learning models: A survey’, Computers and electronics in agriculture 155, 257–282.
5. Jeyalaksshmi, S., Rama, V. & Suseendran, G. (2019), ‘Data mining in soil and plant nutrient management, recent advances and future challanges in organic crops’, International Journal of Recent Technology and Engineering (IJRTE) pp. 213–216.
6. Alizamir, M., Kisi, O., Ahmed, A. N., Mert, C., Fai, C. M., Kim, S., Kim, N. W. & El-Shafie, A. (2020), ‘Advanced machine learning model for better prediction accuracy of soil temperature at different depths’, PLoS One 15(4), e0231055.
7. Suchithra, M. & Pai, M. L. (2020), ‘Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters’, Information processing in Agriculture 7(1), 72–82.
8. Mankar, A. B. & Burange, M. S. (2014), ‘Data mining-an evolutionary view of agriculture’, International journal of Application or Innovation in engineering and management 3(3), 102– 105.
9. Ding, Y., Wang, L., Li, Y. & Li, D. (2018), ‘Model predictive control and its application in agriculture: A review’, Computers and Electronics in Agriculture 151, 104–117.
10. Anitha, A. & Acharjya, D. (2018), ‘Crop suitability prediction in vellore district using rough set on fuzzy approximation space and neural network’, Neural Computing and Applications 30, 3633–3650
11. Jagadeeswari, M., Manikandababu, C., Kavya, S., Prasanna, R. & Madumitha, S. (2022), Artificial intelligence based crop recommendation system, in ‘2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA)’, IEEE, pp. 1127–1133.