# Soil Fertility Detection and Crop Prediction using IoT and Machine Learning

Bhushan Chaudhari[1*], Sachin Kamble[2], Madhuri Patil[3], Gayatri Bhosale[4], Kavita Jagtap[5], Gaurav Patil[6], Priyanka Wakalkar[7]

[1,2,3,4,5,6,7]*Dept. of Information Technology, SVKM's IOT Dhule, Maharashtra, India; Email: chaudharibs@gmail.com*

**Abstract:** India has huge agriculture heritage. This is the major source of livelihood for most Indian families. Farmers are seen to use fertilizers in inappropriate proportion to enhance crop yield which results in infertile land. To overcome this issue, to check the fertility level of the soil, environment conditions and predicting suitable crop and fertilizers required is need of hour. Soil fertility depends on nutrients like Nitrogen (N), Phosphorous (P), and Potassium (K). It is also affected by environmental factors such as temperature, moisture, humidity, etc. The proposed system provides a cost-effective solution using IoT and Machine Learning based approach to check the NPK concentration present in the soil. Based on which, user can predict the soil suitable crop. The technique used comprises an integrated light transmission and detection system which consists of three LEDs with different wavelengths. Photodiode (LDR sensor module) is used for light detection purposes. The output obtained from the photodiode is handled using a Arduino UNO microcontroller. Based on the inputs received from LDR module, NPK concentration can be evaluated. The model is trained with the Crop Prediction dataset to predict the crop using LightGBM algorithm. The proportion of NPK nutrients and the predicted crop is sent to the user as a text message through the GSM module and ThingSpeak cloud platform.

**Keywords:** NPK, Micro-controller, IoT, Machine Learning, LightGBM, LED, ThingSpeak cloud platform, Photodiode.

## 1. INTRODUCTION

The growth of a country's overall economy is largely dependent on the agricultural sector. The population of India increases day by day. India is the second-highest populated country in the world. 70% of rural Indian households get their income mostly from agriculture. In India, 82% of farmers are marginal and small. The total amount of food grains produced in India in 2017–2018 was 275 million tons. [1, 2].

Nowadays crop failures are one of the reasons behind farmers' suicides. [3] Therefore it is necessary to check the soil fertility before taking the crop in the soil. For optimal growth of the crop, land should have adequate fertilizers for good yields and food production. Plants extract nutrients from the soil which are required for their growth. These nutrients are classified into two types:

i. Micronutrients: These nutrients are taken up in large amounts. It includes Chlorine (Cl), Boron (B), Copper (Cu), Iron (Fe), Molybdenum (Mo), Nickel (Ni) and Zinc (Zn).

ii. Macronutrients: These nutrients are taken up in small amounts. It includes Nitrogen (N), Magnesium (Mg), Sulphur (S), Potassium (K), Phosphorus (P) [4].

Out of the macronutrients which are present in the soil, Potassium, Nitrogen and Phosphorous are important nutrients for the growth of plants. While Potassium stimulates flowering and fruiting as well as maintaining nutritional and water balance in a plant cell, Phosphorus stimulates root growth and Nitrogen encourages the growth of leaves and other vegetation [5, 6]. Poor plant production is caused by insufficient fertilizer use, whereas contaminated soil is caused by excessive fertilizer use [7]. Therefore, for good plant growth, these components are crucial and must be present in the soil in the proper quantity.

The majority of farmers spray fertilizer into the soil directly due to a lack of infrastructure. Because of financial limitations and the unavailability of a nearby soil testing facility, farmers cannot regularly measure the nutrient content of the soil. Furthermore, it takes a long time for farmers to get the findings of soil tests. There is need of a method that is both efficient and portable to measure the nutrient levels in soil, as it is important to have a convenient method to accurately measure the soil nutrient levels in various locations [5, 6].

Various methods have been used by previous authors, including optical techniques, electrical, and electromagnetic techniques, electrochemical techniques, mechanical, acoustic, chemical, ICP and fluorescence spectroscopy approaches. Nearly each of these techniques, however, areone of the other prohibitively expensive, call for intricate setups, or is ineffective for determining nutrients [7].

The proposed system is built using methods that take these concerns into account and makes use of IoT and machine learning techniques to reduce costs and deliver outcomes in the shortest period.

By considering this issue, the proposed system is done together using IoT and Machine Learning techniques. It can reduce costs and give results to the user in a minimum amount of time. The purpose of this research is to create a system that can measure the proportion of nutrients in the soil, predicted crops that are suitable for the soil with more accuracy, and maximizes the crop yield. NPK nutrients in the soil have distinctive optical properties, therefore the detection system is based on the absorbing principles to identify NPK. Different wavelength LEDs were employed in the light transmission system, soil interacts with the light by absorbing it, and a photo-diode that converts light into electrical energy detects the remaining light. The rate of absorption is calculated from the photodiode's output and the output current is translated into voltages using a microcontroller. based on this data, NPK concentration gets evaluated and the respective message is sent to the user.

## 2. RELATED WORK

Various sensors were used to measure the temperature, pH, humidity, moisture, and nutrient levels. Monitoring these ensures that, the soil is fertile enough depending on the nutrients present. Farmers receive the information via GSM and are shown it. Therefore, it reduces the amount of waste generated from husk growth, resulting in higher-quality and healthier crops [8,19]. As a result of its ability to identify the nutrients present in the dry soil samples without any need for laborious sample preparation techniques, optical soil nutrient detection techniques are used to create portable sensors. The research examines both practical and successful spectroscopic procedures, as well as technologically innovative and cutting-edge methods such as imaging systems, microfluidics, and MEMs-based sensors [7].

A soil fertility test kit and capsules are used to find the levels of Phosphorus, Potassium, and Nitrogen present in the soil. This method is made up of three different types of capsules that test for NPK. A soil sample is taken and mixed with special capsules that contain chemicals. This mixture is shaken vigorously, and the resulting color change is observed. This color change is then compared to a set of pre-determined colors that represent different levels of nutrients in the soil. By matching the color of the mixture to the pre- determined colors, the test can determine how much nitrogen, potassium, and phosphorus is needed in the soil to ensure healthy plant growth [1]. The soil monitoring system that makes use of wireless sensing networks has made it feasible for the Thing Speak cloud platform to remotely monitor moisture levels, temperature, humidity, and NPK values. A mobile app and a smart plan can help farmers to improve the fertility of their soil and increase crop yields [9].

To identify the levels of (N, P, K) and other nutrients inside the soil, an electrochemical sensor has been designed. The sensor operates under the ion absorption from soil aqueous solution theory. The sensor and Arduino Microcontroller circuits are designed to identify the inadequate soil component for analysis [10]. A special sensor using fiber optics was used to measure the levels of potassium, nitrogen, phosphorus in a soil sample. The sensor works by detecting the color of the soil after certain chemicals are added, which can indicate the presence of these elements. The sensor can classify the levels of these elements as low, high, medium, or none. It is used to identify which elements are lacking in the soil [11]. An improved genetic algorithm (IGA) that suggests different crop settings using time-series sensor data. The handling of exploitation and exploration for adjusting the parameters to achieve the maximum yield is then given using a neighborhood-based technique. By offering the nutrient estimation of ideal conditions, the approach is used to enhance soil fertility performance [12]. The system comprises Bluetooth to enable data transfer to adjacent mobile devices. A pH electrode composed of antimony is used. The DS18B20 sensor, which operates on the Dallas one-wire protocol, is used to measure soil temperature. The STM32 Nucleo platform is used to develop the entire system [13].

A model that can aid farmers in forecasting a variety of soil properties, such as calcium Ca, phosphorus P, and pH values are created using the machine learning (ML) method. According to the measured Ca, P, and pH levels, these features may help farmers decide how much fertilizer to put into the soil sample to maintain consistent soil fertility [15, 18, 21]. To increase prediction accuracy, a variety of machine learning techniques, including the Linear Regression algorithm, Artificial Neural Network, Gaussian Naive Bayes, Random Forest method, Gradient Boosting method, K- Nearest Neighbors algorithm, Decision Tree algorithm, and Logistic regression, were tested [23].

Colorimetry is a method used to measure the color of a sample, in this case, a soil sample. By using a specific liquid to extract the soil, we can create a solution that represents the color of the soil. This solution is then measured using a color sensor, which gives us a numerical value that represents the color of the soil. This value is then compared to a reference standard and can be used to classify the soil as having low, medium, or high levels of certain elements or nutrients that affect the soil color. It is also used to classify the soil by using Naive Bayes classification algorithm. It predicts the accuracy of the targeted system after using the Naive Bayes classifier, XGBoost, Support vector machine. Thus, the desired approach is advantageous in terms of reducing the time necessary for evaluating soil fertility and assessing the correctness of our results [22].

A crop recommendation system based on fuzzy logic to help rural farmers. The proposed model is intended to deal with eight important crops grown in West Bengal. To facilitate faster parallel processing, separate fuzzy rule bases were constructed for each crop. A diversified dataset was used to validate the model's performance [25].

## 3. ARCHITECTURE AND METHODOLOGY

By considering all the possibilities of the solution, here the methodology is proposed to detect NPK nutrients. The proposed methodology is outlined in a diagram, shown in Figure 1.

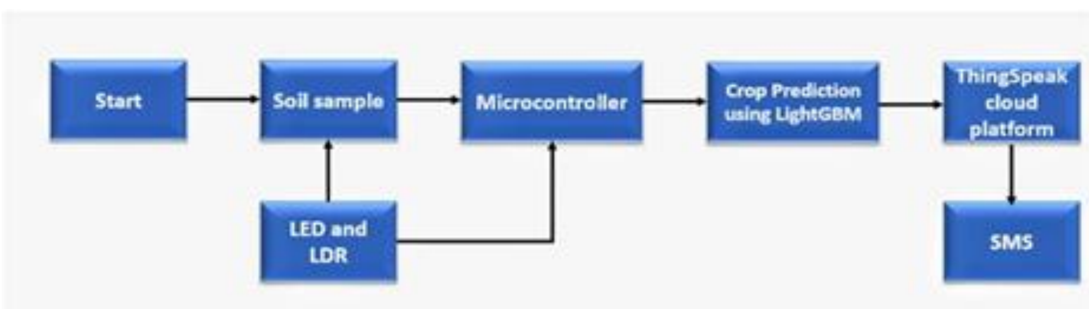


**Fig 1.** Block Diagram of the proposed system

The proposed method is a system that uses three different components to measure the properties of soil. The first component is a light sensor, which is used to measure the amount of light that passes through the soil sample. The second component is a cloud platform, which is used to collect, process and store the data from the sensor in the cloud. The third component is a prediction algorithm, which analyses the data from the sensor and the cloud platform and makes predictions about the soil properties such as its nutrient levels. All of these components work together to provide a comprehensive and efficient way of measuring the soil. It uses an Arduino UNO microcontroller to operate the source. Under the LED light's illumination, the soil sample is positioned on top of the reflector at the optimal thickness. Beer's Law states that the equation for absorbance (Ab) is as follows:

$$Ab = -log_{10}\ (T1/I1) \qquad (1)$$

Where,

Ab: Absorbance,

T1: transmitted light,

I1: incident light.

To measure the path length LED, photo-diode, and reflectorare in the proper position as shown in Fig 2.

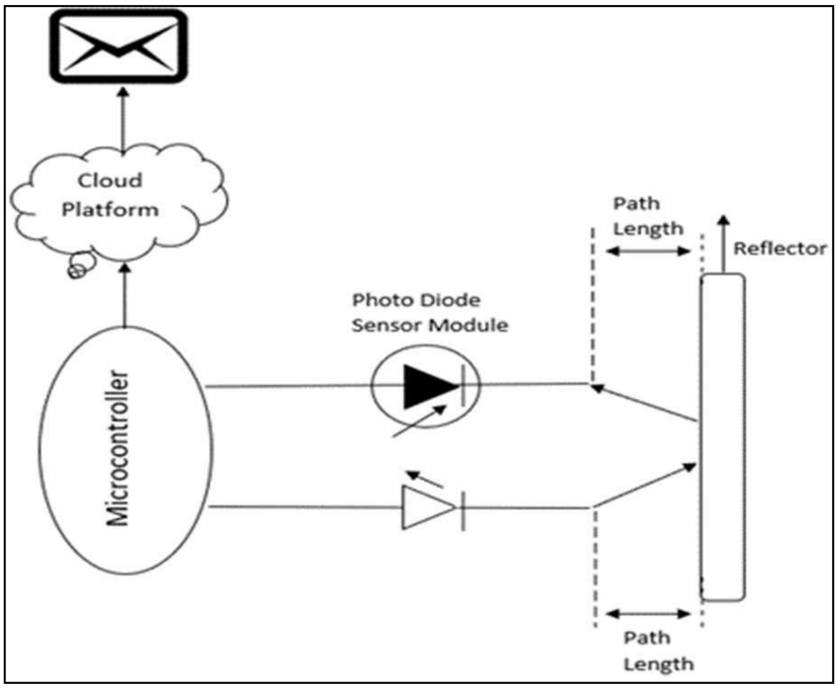


**Fig 2.** The architecture of the experimental setup

Three LEDs, each with a different wavelength, are used in the light transmission system as shown in table 1. Each LED is selected based on the wavelength of the spectrum that the NPK soil absorbs. These LEDs are placed one by one over the soil and positioned parallel to the LDR sensor module. Here, LDR is used as a photodiode in a light detection system to capture the light from LED. It is a type of semiconductor which converts light into current. Soil absorbs the light from the LED as molecules of the soil have a relation with the electrons and the remaining light gets reflected which isdetected by the photo-diode and the absorption rate is measured. The equation for the conversion of light into the current is as follows:

$$I = \frac{R_S}{R_S + R_L} * V \qquad (2)$$

Where,

I: current,
V: voltage,
Rs: Resistance,
$R_L$: Photocell resistance.

Standard values of wavelength and the absorptionwavelength (nm) of NPK are given below:

**Table 1: Standard values of LED**

| LED Type | Nutrient | Wavelength(nm) | Absorption Wavelength(nm) |
|---|---|---|---|
| LED1 | Nitrogen(N) | 460 - 485 | 390 - 410 |
| LED2 | Phosphorus(P) | 500 - 574 | 470 - 500 |
| LED3 | Potassium(K) | 635 - 660 | 530 - 540 |

Photo-diode converts light intensity into voltages and then the data is sent to the Microcontroller for digital conversion and further processing. The output of the photodiode is the percentage of the NPK nutrients. This percentage proportion of NPK nutrients gets compared with the dataset values and the corresponding suitable crop is predicted. Dataset has 2200 rows and 11 columns (N, P, K, N_per, P_per, K_per, Temperature, Humidity, rainfall, and label). In this dataset, the proposed system deals with the N_per, P_per, and K_per attributes as independent variables and 'label' as a dependent variable. This dependent variable contains 22 crops. Fig 3. Illustrates the required proportion of NPK to each crop.

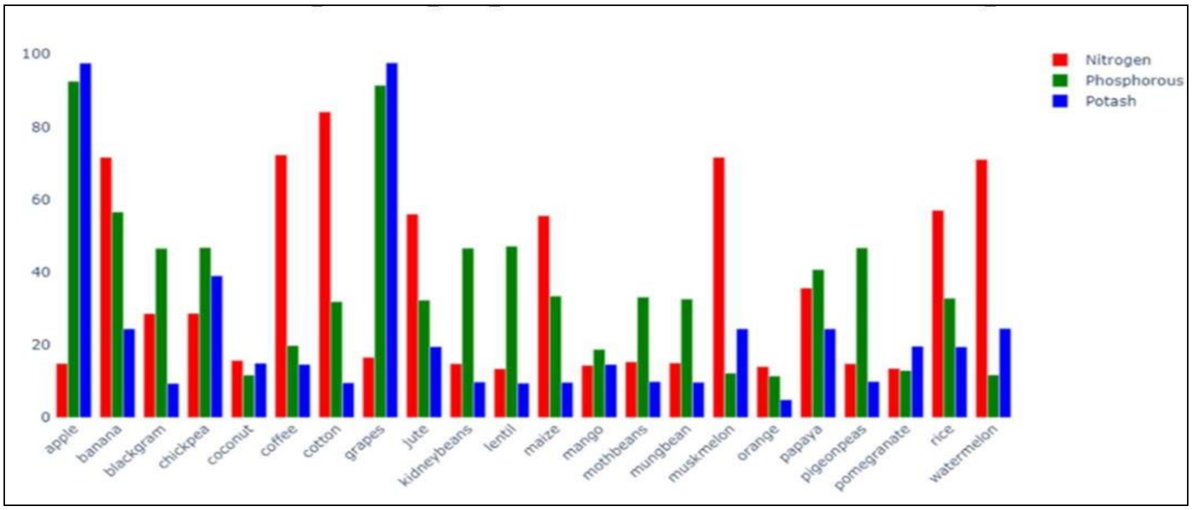


**Fig 3.** Description of NPK nutrients to each crop

This system utilizes the ThingSpeak Cloud platform and GSM module for sending SMS which contains the NPK proportion and predicted crop to the respective user. ThingSpeak establishes a communication connection between the Arduino UNO microcontroller and the cloud platform. The Wi-Fi Module connects the board to this channel. To connect to the cloud, API keys are used. The channel readapikey and writeapikey have two API keys.

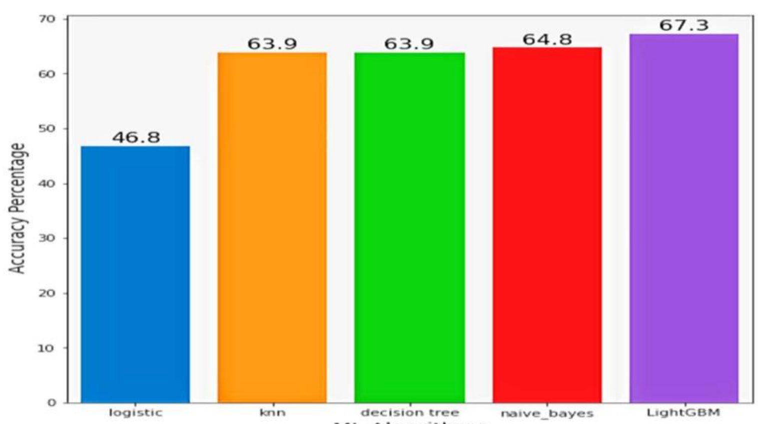


**Fig 4.** Accuracy comparison of different algorithms

Here LightGBM algorithm is used for training, testing, and prediction purpose because it gives the highest accuracy than other algorithms as shown in Fig 4.

## 4. COMPARATIVE ANALYSIS

The laboratory evaluation method is widely utilized, but findings take months to obtain because it utilizes human effort to collect data and transport it to the laboratory; also, in most rural locations, lab facilities are not available to neighboring fields [2,4,6]. The Chemical method is rarely used because it is time-consuming lengthy and costly, also sometimes it is harmful to human beings [1,5]. The electrochemical method and fiber-optic methods require more time and are costly, therefore it is rarely used [4,14]. Building a low-cost, portable soil nutrient sensor can use visible-infrared (Vis-IR), inductively coupled plasma (ICP), fluorescence, and colorimeters are comes under the optical methods, but these methods are impacted by ambient conditions that reduce the accuracy of sensor results [7,12]. IoT technology is being used in the agricultural sector to help farmers solve important problems by collecting and analyzing data from IoT-enabled devices, such as sensors. [9]. To determine soil fertility, machine-learning techniques including Linear Regression ML Algorithm, Random-Forest- algorithm, Gradient Boosting method, K - Nearest Neighbours algorithm, Support Vector Machine, Decision Tree algorithm, and Artificial Neural Network are utilized [17, 20]. The machine-learning method increases the precision of predicting soil parameters [15, 16]. Therefore, by together using IoT and ML we can reduce costs and give rapid results to the user.

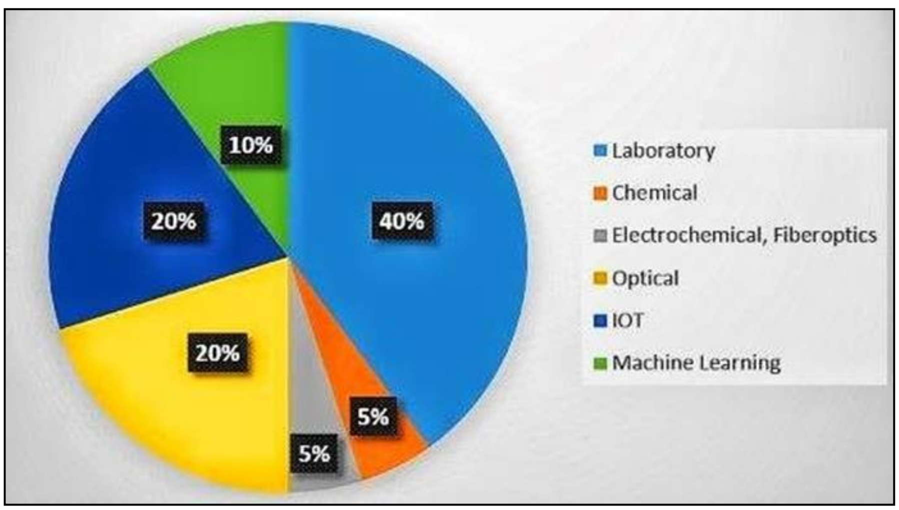


**Fig 5.** The proportion of methods used till now for NPK detection

## 5. RESULTS AND DISCUSSION

The proposed system uses the LightGBM algorithm which is used to predict the crop. It helps to improve the model efficiency and reduces memory usage. It is also a gradient- boosting framework based on the decision tree. The dataset used for the prediction is taken from Kaggle that included different parameters which are required for better crop yield. According to the model, the prediction accuracy of the LightGBM algorithm is 67.33%. The system evaluates the NPK proportion with the help of a microcontroller then the data is compared with the dataset values and a suitable crop is predicted successfully as shown in Fig 5.

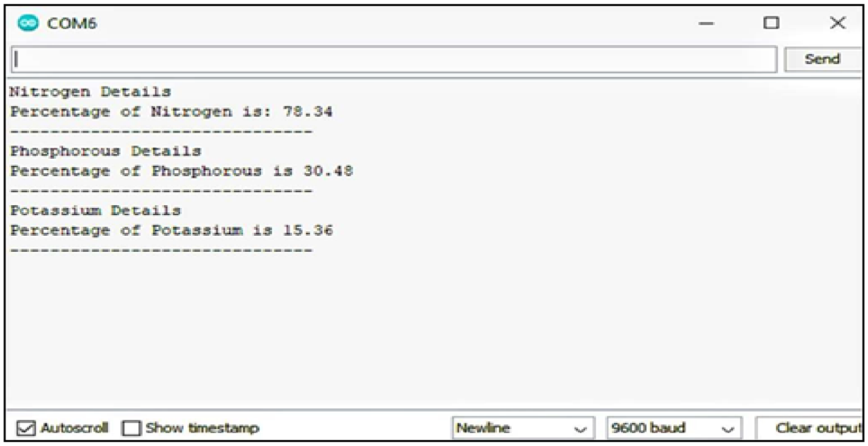


**Fig 6.** Percentage of NPK nutrients present in the soil

The percentage of NPK nutrients shown in Fig. 5 are compared with the dataset using the LightGBM algorithm. Considering the comparison, the crop is predicted as depicted below.

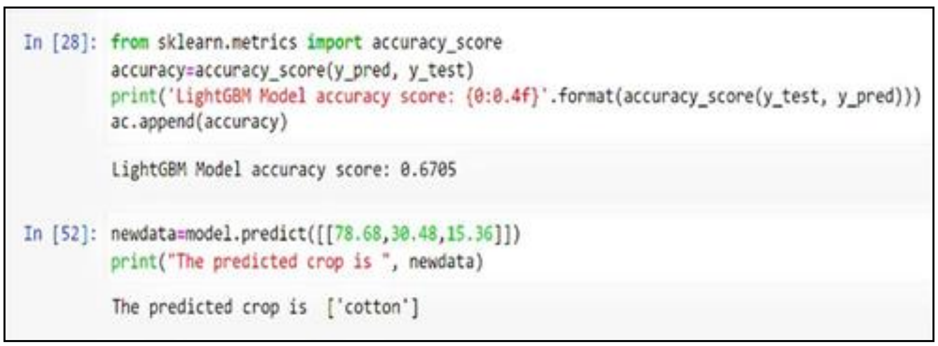


**Fig 7.** Predicted crop based on the NPK percentage

## CONCLUSIONS

NPK promotes plant growth most effectively and the deficiency of these nutrients causes the plant to lose its yields. The traditional method for NPK detection is complex and time-consuming. IoT and Machine Learning technology are very helpful for farmers for getting updated information. The combination of IoT technology to collect data and Machine Learning algorithms to analyse the data, allows farmers to optimize their farming practices, resulting in more efficient use of resources and higher crop yields. By knowing the accurate values of the NPK nutrient, farmers can grow a suitable crop and apply the required fertilizer to the soil. Therefore, in this paper, the proposed methodology focuses on providing accurate information to the farmers regarding soil nutrients and the predicted crop which is suitable for soil in the form of SMS using a cloud platform. The proposed method will be beneficial for government entities like municipal corporations, etc. By determining the soil's health, this system will be useful for predicting barren land. The analyzed data from the Machine learning model will be saved and used for future reference. It is less expensive than other technologies and can reduce problems with determining the number of nutrients available in the soil. It will boost the yield of quality crops and decrease the overuse of fertilizers and increase good crop production. Including a GPS in the GSM module, the system will predict the temperature, humidity, and moisture by fetching the respective user's location. By extending these fertility measurement parameters the system will give more accurate results.

## REFERENCES

[1] H. Pallevada, S. Parvathi, T. V. Munnangi, B. Rayapudi, S. Gadde, and M. Chinta, "Real-Time Soil Nutrient detection and Analysis," *ICACITE 2021*, Mar. 2021, pp. 1035–1038.

[2] Lydiah Gatere , Dimkpa, and P. Bindraban, "Methods for Rapid Testing of Plant and Soil Nutrients", 4th International Conference IEEE, 2017

[3] P. Raju Iyer, "Indian Farmers' Suicides and Interventions to Combat the Problem." (2021). Google Scholar

[4] S. S. Patil and V. A. Shukre, "Comparative Study of Different Methodologies used for Measuring Soil Parameters," (IJERT), ISSN: 2278-0181, Conference Proceedings 2020.

[5] A. Amrutha, A. Sreedevi, and R. Lekha, "Fertilizer dispensary system and automatic soil nutrient detection," CTFC, IEEE 2016

[6] M. Masrie, R. Sam, and S. Alam, "Detection of Nitrogen, Phosphorus, and Potassium (NPK) nutrients of soil using Optical Transducer," Fourth IEEE ICSIMA, 2017.

[7] M. M. Shirolkar, P. S. More, A. J. Verma, ": Mini review: Determination of soil nutrients (NPK) using optical methods," vol. 44, no. 12. Bellwether Publishing, Ltd., 2021

[8] G. Yoshitha, A. V. Mutyalamma, and A. Dakshyani, "IoT based Smart Agriculture to Measure soil fertility parameters (Temperature, Ph, Humidity, Moisture) and Nutrient Values of the Soil," IJEAT, vol. 9, no. 5, Jun. 2020.

[9] S. Annigeri, V. Bagali, and S. Nuchhi "IOT Based Soil Testing Instrument for Agriculture Purpose," 1st Bangalore Humanitarian Technology Conference, IEEE, Oct. 2020.

[10] B. Krithiga and R. Sindhuja "Soil Nutrient Identification Using Arduino," Volume 1 Issue 4, Asian Journal of Applied Science and Technology, May 2017.

[11] M. Bhoi and L. Gavade "N, P, K Detection & Control for Agriculture Applications using PIC Controller: A Review," IJERT, Vol. 6 Issue 04, ISSN: 2278 – 0181, April-2017

[12] U. Ahmed, G. Srivastava, J. Lin, "Evolutionary computation based nutrient recommendation system for soil fertilization," vol. 189, Oct. 2021.

[13] A. Na, S. Varshney, and W. Isaac, "An IoT based system for remote monitoring of soil characteristics," INCITE - The Next Generation IT Summit on - IOT, Feb. 2017.

[14] L. Gavade, "N, P, K Detection using Fiber Optic Sensor and PIC Controller," IJESC, Volume 7, July 2017.

[15] N. Janvier, N. Arcade, N. Eric, and N. Jean, "Machine Learning based Soil Fertility Prediction," IJISET, ISSN (Online) 2348 – 7968, Vol. 8 Issue 7, July 2021.

[16] N. Bhangale, S. Agarwal, S. Gavhane, and K. Dhanure, "Application of Colorimetry to Determine Soil Fertility through Naive Bayes Classification Algorithm," Ninth ICCCNT, IEEE – 43488, July 2018.

[17] O. Chambers, J. Trontelj, "Strategy for soil nutrients prediction using spectroscopic method," *Sensors*, vol. 21, MDPI 2021

[18] T. L. Surekha, Sai Teja, and P. Gowtham, "Soil Analysis for Fertilizer Prediction using IoT and Machine Learning," IJCA, ISSN: 2005-4297 IJCA, Vol 13, No.2, pp. 1497-1504, 2020.

[19] M. Khakal, M. N. M. Deshpande, and M. Varpe, "Measurement of NPK from PH value, e-ISSN No.:2349-9745, IJMTER 2016.

[20] Javier Lacasta, Borja Espejo-García, F. Lopez-Pellicer, Javier Nogueras-Iso, "Agricultural recommendation system for crop protection", Volume 152, 2018, ISSN 0168-1699.

[21] G. Banerjee, I. Ghosh and U. Sarkar, "A Fuzzy Logic-Based Crop Recommendation System", Proceedings of ICFCS, vol 1255. Springer, 2021.

[22] J. Lachure, G. Dhabal and R. Doriya, "Crop Recommendation System with Cloud Computing," Third ICIRCA, IEEE, 2021, pp. 1404-1411.

[23] J. Haban, E. Sybingco, A. Bandala and E. Dadios, "Fuzzy Logic based Soil Fertilizer Recommendation System," IEEE 2020, pp. 1171-1175.

DOI: https://doi.org/10.15379/ijmst.v10i2.2855