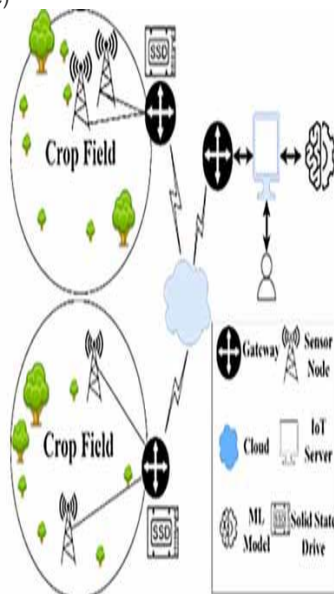


### Abstract:

An accurate amount of fertilizer according to the real-time context is the basis of precision agriculture in terms of sustainability and profitability. Many fertilizers recommendation systems are proposed without considering the real-time context in terms of soil fertility level, crop type, and soil type. The major obstacle in developing the real-time context-aware fertilizer recommendation system is related to the complexity associated with the real-time mapping of soil fertility. Furthermore, the existing methods of determining the real-time soil fertility levels for the recommendation of fertilizer are costly, time-consuming, and laborious. Therefore, to tackle this issue, we propose a machine learning-based fertilizer recommendation methodology according to the real-time soil fertility context captured through the Internet of Things (IoT) assisted soil fertility mapping to improve the accuracy of the fertilizer recommendation system. For real-time soil fertility mapping, an IoT architecture is also proposed to support context-aware fertilizer recommendations. The proposed solution is practically implemented in real crop fields to assess the accuracies of IoT-assisted fertility mapping. The accuracy of IoT-assisted fertility mapping is assessed by comparing the proposed solution with the standard soil chemical analysis method in terms of observing Nitrogen (N), Phosphorous (P), and Potassium (K). The results reveal that the observations by both methods are in line with a mean difference of 0.34, 0.36, and -0.13 for N, P, and K observations, respectively. The context-aware fertilizer recommendation is implemented with the Logistic Regression (LR), Support Vector Machine (SVM), Gaussian Naïve Bayes (GNB), and K-Nearest Neighbor (KNN) machine learning models to assess the performance of these machine learning models. The evaluation of the proposed solution reveals that the GNB model is more accurate as compared to the machine learning models evaluated, with accuraci... (Show More)



The major obstacle in developing the real-time context-aware fertilizer recommendation system is related to the complexity associated with the real-time mapping of soil CCBY - IEEE is not the copyright holder of this material. Please follow the instructions via <https://creativecommons.org/licenses/by/4.0/> to obtain full-text articles and stipulations in the API documentation.

### SECTION I.

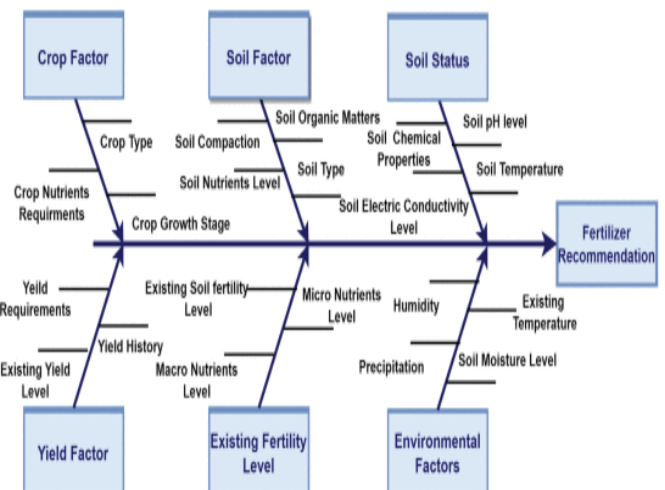
#### Introduction

The increase in human population coupled with the decrease in natural resources has raised concerns for food

security [1]. Efforts were made to improve productivity in agriculture to feed the ever-increasing human population [2]. Extensive use of fertilizers to improve productivity is the core of these efforts that results in inefficient use of resources. The injudicious use of fertilizers to improve productivity has further diversified the issues related to soil deterioration and sustainable development [3]. Every effort was made to improve the productivity in agriculture in the last decades. Precision and smart agriculture are evolving to improve productivity while maintaining yield [4]. Efficient use of resources while improving productivity is the core of sustainable developments in agriculture through smart and precision agriculture [5]. There is an immense need for efficient fertilizer application according to context to using smart agriculture practices to support sustainable developments in agriculture.

Precision and smart agriculture have emerged as a new paradigm with the emergence of new technologies for better productivity in agriculture with the optimal usage of resources [6], [7]. The management of soil fertility is the core of sustainable agriculture to produce food for the ever-increasing human population [8]. Site-specific fertilizer applications have environmental, economic, and yield advantages [9]. Internet of Things (IoT) and machine learning potential technologies for effective soil fertility management.

Soil fertility is the level of nutrients present in the soil that are essential for plant growth. The growth of plants is directly linked with the soil fertility level. Extensive crop production has resulted in the depletion of soil nutrients level resulting in low soil fertility. The extensive agriculture activities and the injudicious use of fertilizer to replenish soil fertility have raised the issues of soil deterioration in terms of soil salinity, soil aridity, and soil acidity. The soil fertility problems are directly linked to the injudicious usage of fertilizers without any recommendations to improve yield and productivity in agriculture. The selection of appropriate fertilizer is based on multiple criteria. Factors affecting the selection of appropriate fertilizer are shown in Figure 1.



**FIGURE 1.** Factors affecting fertilizer recommendation.

[Show All](#)

The efficient application of fertilizers can save resources as well as prevents soil deterioration. The inadequate

nutrient management raises many concerns regarding low productivity, and soil deterioration [10]. The efficient use of accurate fertilizer according to the context is the basic requirement for better crop yield, productivity, and sustainable soil quality. However, it is quite challenging to fulfill these requirements. The major issue regarding the application of precise fertilizer is the complexity associated with the traditional method of determining soil fertility.

To overcome the difficulties associated with the traditional method of soil fertility mapping, an IoT-assisted soil fertility mapping architecture is also proposed and implemented. The traditional method of soil fertility observations is the major obstacle in the implementation of precision fertilizer recommendations according to the context. Soil fertility is determined using soil chemical analysis to ascertain the existing level of soil nutrients. The standard method of soil chemical analysis for fertility level determination is costly, time-consuming, and complex. There is a need for a solution for the assessment of existing fertility levels that are easy to use by farmers for smart agriculture applications. With the emergence of new sensing technologies, it is possible to overcome the problems associated with the existing soil fertility mapping and assessment methods [11].

The emergence of new sensing capabilities and revolutionary IoT and machine learning technologies are ideal technologies for effective soil fertility management to overcome the problems associated with traditional fertility mapping and monitoring [12]. IoT and machine learning are revolutionary technologies that have shown revolutionary changes in every sphere including agriculture in the form of smart agriculture [13], [14]. The directly sensed crop field fertility enables efficient and effective management of soil fertility.

For accurate soil fertility management and appropriate fertilizer recommendation, we proposed an IoT-assisted soil fertility mapping architecture to assist the machine learning-based context-aware fertilizer recommendations according to the real-time soil fertility level, crop type, and soil type. The major contributions of the study are listed below

### A. Contribution of the Study

The unique contributions of the study are as follows:

1. We propose an IoT-assisted fertility mapping architecture to overcome the problems associated with the traditional soil fertility assessment method.
2. The study assesses the accuracy of proposed IoT-assisted fertility mapping against the standard soil chemical analysis method.
3. The study proposes a precise fertilizer recommendation system based on the real-time sensed soil fertility level, soil type, and crop type using machine learning models.
4. The study compares the performance of the Logistic Regression (LR), SVM, KNN, and GNB models used for the fertilizer recommendation system.

### B. Organization of the Study

The study is organized into the following sections: Section II explores the recent developments regarding soil characteristics mapping and fertilizer recommendations to identify the research gap for the study. Section III elaborates on the material, proposed IoT architecture, machine learning model, dataset, and the implementation details of the proposed solution. Section IV unveils the experimental results. We conclude and summarize the study in Section V.

## SECTION II.

### Literature Review

Machine learning and IoT are extensively applied in agriculture for smart farms and precision agriculture applications. The emergence of new sensing, communication, and processing capabilities has created the potential for smart agriculture with exciting new services and applications. IoT-based crop field monitoring is quite common for different purposes with different sensing capabilities, such as IoT-assisted farm monitoring and control.

Q. V. Khanh et al. [15] explored the role of 5G in the development of IoT along with the vision, applications, and challenges for 5G in IoT applications. Khanh Quy et al. [5] reviewed different IoT enabler technologies for IoT smart agriculture. The study discussed the vision of IoT technologies for smart agriculture applications. Johannes Tiisanen M. [16] proposed soil Scouts named underground soil sensor node architecture to receive sensed data from one kilometer and maintenance-free sensor node. Boursianis A. et al. [17]. Boursianis A. et al. [18] explored the communication, sensing technologies, and communication mechanisms for the Internet of Underground things (IOUT) for precision agriculture. The study explores the possibilities of soil characteristics sensing using the IoT. Akhter F. et al. [19] discovered the role and potential of IoT, machine learning, and data analytics technologies in agriculture. Apple disease prediction model is also proposed using data analytics, machine learning, and IoT.

In [25], the authors use IoT and machine learning for smart irrigation water and reference Evapotranspiration (ET) determination according to the prevailing environmental conditions [20]. Zhiming Hu et al. [21] propose an IoT and machine learning-based reference evapotranspiration (ET<sub>o</sub>) system for precision irrigation water applications. Indra et al. [22] suggest an agriculture recommendation system for Indian farmers. Ponnusamy and Natarajan [23] illustrate the use of modern technologies for smart farming. Michele E. et al. [24] evaluate smart technologies to determine their impact on agriculture in terms of efficient use of resources, sustainability, and cost reduction. The study also explores the impacts of IoT, Artificial Intelligence (AI), and data analytics in precision agriculture. Sinha B. and Dhanalakshmi [4] explore the major IoT components, challenges, and the future of IoT applications in agriculture. Rehman A. et al. [25] review various IoT technologies for smart monitoring and control in agriculture. Parween S. et al. [26] evaluate smart agriculture using IoT for improvements in productivity in agriculture.

Liu Z. et al. [27] propose an IoT and machine learning model for the prediction of plant diseases based on the correlation between plant disease and environmental conditions. Suresh G. et. [28] propose a machine learning model for the estimation of crop yield. Choudhary M. et al. [29] propose a machine learning-based plant disease

prediction and crop recommendation system. Madhuri and M Indiramma [30] propose an Artificial Neural Network (ANN) based crop recommendation system according to the climatic conditions, soil type, and crop characteristics. The proposed crop recommendation is a promising aspect of crop planning with an accuracy of 96% in predicting the crop type.

Apart from smart irrigation and disease prediction, IoT and machine learning are extensively used for the mapping of soil characteristics. Deshmukh M. et al. [31] propose a farming assistant model for soil fertility improvement using crop prediction with the application of the XGBoost model. Gosai D. et al. [32] propose a Random Forests (RF) based crop type recommendation system based on IoT-assisted soil characteristics mapping. The IoT aspect of the soil characteristics mapping is performed by observing the soil's major nutrients NPK, soil pH, soil temperature, and soil moisture. The proposed model uses the XGBoost model for crop type recommendation with 99% accuracy. In terms of recommendations, MobileNet is recommended for plant disease identification with the help of plant images. They apply RF for the fertilizer recommendation according to prevailing soil fertility.

Pruthviraj et al. [33] propose a machine learning-based soil classification and recommendation based on the soil fertility level and crop type. The study also compares the performance of SVM, decision tree, and KNN for the fertilizer recommendation systems in terms of accuracy. The results reveal that SVM is the most accurate machine learning model for fertilizer recommendation according to soil classification. Rab Nawaz et al. propose an IoT-based salinity mapping at irrigated scheme levels [34].

Swaminathan B et al. [35] propose an IoT and machine learning-driven fertilizer recommendation system using deep neural networks based on expert opinion. Quaye et al. [36] propose integrated fertility management techniques for farmers in Ghana. Mesfin S. et al. [37] propose a site-specific fertilizer recommendation system for improving barley crops based on soil nutrient supplies. Amgain L. [38] propose a fertilizer recommendation system for rice crops to improve profitability, productivity, and fertilizer usage efficiency. JuhiReshma S R et al. [39] propose the estimation of fertilizer quantity according to crop requirement and soil fertility.

Firmansyah [40] propose an AI-based expert system for fertilizer recommendations in palm oil crops. Gorai T. [10] discusses the RS-based approaches for soil fertility assessment in terms of effectively utilizing fertilizers using different indices. Jayashree D. et al. [41] propose a fertilizer and pesticide recommendation system based on the tree data structure. The proposed solution is based on the YOLO algorithm in conjunction with Convolution Neural Network (CNN) model. Erick Firmansyah et al. [40] recommend an AI-based expert system for the accurate utilization of fertilizer. Udoumoh U. and Ikrang [42] review new technologies, such as GIS, GPS, and Remote Sensing (RS) for soil nutrient management to develop an efficient fertilizer recommendation system.

The soil nutrients are assessed by colorimetry and machine learning model with Gaussian Process Regression (GPR). Shweta Singh et al. [43] propose multi-criteria decision-making techniques for an appropriate fertilizer recommendation system. The proposed solution uses the AHP method and TOPSIS metacriterion techniques for comparing and evaluating the choices. The selected

fertilizer choices for the fertilizer recommendation are Urea, Di-Ammonium Phosphate (DAP), and Potassium Chloride. D. Hassan and M. Manohar [44] propose IoT-based site-specific soil nutrients management systems, such as Nitrogen (N), Phosphorus (P), Potassium (K), and pH of the soil.

Koli S. et al. [45] propose an intelligent soil wetness system using IoT to accurately analyze soil moisture and nutrient analysis. Ahmed U. et al. [46] review the fusion sensing technologies and their integration on different platforms for crop parameter monitoring like nitrogen, chlorophyll level, and leaf area index. Andrianto H. et al. [47] propose and evaluate the service system platform using IoT to assess the plant's nutritional deficiencies and fertilizer recommendation according to the status of soil nutrients level. Nutini F. et al. [48] propose an imagery approach for the soil Nitrogen (N) level assessment using crop modeling, sentinel imagery, and the soil nitrogen level index.

J. Singh and V. Singh [49] propose the efficient usage of Nitrogen (N) with the help of crop reflectance to assess the Nitrogen (N) level and site-specific fertilizer management in cereal crops. Kassa M. et al. [50] propose a Quantitative Evaluation of the Fertility of Tropical Soil (QUEFTS) model to assess the relationship between the site-specific fertilizer applications and yield attained in barley crops. The results from the implementation of the study show that the proposed QUEFTS model can be successfully used to assess the nutrient requirements of crops along with the site-specific fertilizer recommendations.

Wen G. et al. [51] conduct a study to predict crop yield by using a machine-learning approach with the help of a Random Forest Regressor (RFR) model. The study also recommends site-specific fertilizer recommendations using the PFR model based on weather conditions, plant growth, soil characteristics, and spectral index data. MacCarthy D. et al. [52] review different Decision Support Tools (DST) for the efficient use of fertilizer. Cuong N. et al. [53] propose deep learning-based fertilizer and pesticide recommendation systems for precision agriculture.

Many IoT-based solutions for soil characteristics are proposed, while machine learning is applied for decision-making. The real-time context-aware fertilizer recommendation system is not targeted by the existing literature. There is also a need for fertilizer recommendations according to real-time context.

### SECTION III.

#### Materials and Methods

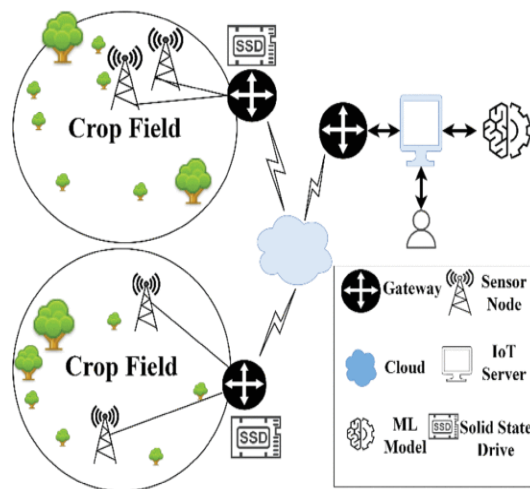
This section discusses the proposed architecture for the IoT-based portable fertility mapping, machine learning model, and dataset for the prediction of fertilizer based on fertility, soil type, and crop type.

#### A. Internet of Things (IOT) Architecture of Soil Fertility Mapping

The IoT architecture enables portable fertility mapping with centralized data storage and decision-making for different stakeholders. The proposed architecture is based on sensing soil fertility mapping in terms of NPK soil macronutrients by using modern sensors. The sensor node is portable; therefore, we can use a single node at multiple sites for fertility mapping. The architecture is. The proposed IoT-based fertility mapping architecture is



simple and lightweight to easily move across the field is shown in. Figure 2.



**FIGURE 2.** Internet of Things (IoT) architecture for soil fertility mapping.

[Show All](#)

The data received at the gateway module is transferred to the server through the cloud. The gateway module can also store data temporarily if cloud connectivity is not available in the field. Upon the availability of Internet connectivity, the temporary data stored at the gateway node is transferred to the server through the cloud. This architecture facilitates fertility mapping in remote areas where Internet connectivity is not available.

The sensed data from the sensor is transferred through Radio Frequency-433 (RF-433) MHz module to the gateway node from where the data is transferred to the server through the Internet. RF-433 is a lightweight, cost-effective communication module used to achieve the objectives of the proposed solution. RF-433 module uses Amplitude Shift Keying (ASK) for transfer data, where the amplitude of the carrier wave is changed in response to the new data. The NPK sensor node and the sampling of soil fertility with the NPK sensor node are shown in Figure 3, and Figure 4, respectively.



**FIGURE 3.** NPK sensor node.

[Show All](#)



**FIGURE 4.** NPK sensor node sampling.

[Show All](#)

Soil fertility is expressed by the presence of nutrients in the soil. The fertility is observed by the presence of three macronutrients (Nitrogen (N), Phosphorus (P), and Potassium (K)) with the implementation of the proposed IoT-assisted fertility mapping architecture. The sensor node is used to observe the NPK nutrient values from the soil. Soil fertility can be observed as per requirements without specialized skills in a cost-effective manner. For

experiment purposes, the soil fertility is observed at twenty sampling points in an area of one acre. To evaluate the accuracy of the proposed IoT fertility mapping the observations are compared against the standard method of a soil fertility assessment. For comparison purposes, the Bald-Altman difference plot is used. The bald-Altman difference plot is used to compare the difference between the observation of the same event by two different methods.

B. Machine Learning Model for Fertilizer Recommendation

The objective of the proposed solution is to recommend an appropriate amount of fertilizer according to the level of soil fertility in terms of macro-elements (NPK), crop type, and soil type. The existing soil fertility is mapped to determine the level of NPK in the soil. The real-time soil NPK level in terms of fertility, soil type, and crop type is used in the machine-learning model to recommend an appropriate amount of fertilizer.

For Implementation, we use Logistic Regression (LR), Gaussian Naïve Bayes (GNB), Support Vector Machine (SVM), and k-Nearest Neighbor (KNN) based machine learning models to predict the fertilizer according to the context. LR is used for classification and predictive analysis in the case of a linear relationship. LR is a supervised machine learning algorithm based on the probability of occurrence of an event from a set of input conditions also known as input features or independent variables. GNB is a classification model based on conditional probability. GNG is suitable when the input feature set is independent. SVM is a supervised machine learning algorithm used for binary and multiple classifications. SVM is suitable for problems with the multidimensional dataset. KNN is a supervised machine learning that groups similar data nearby. The new data is put into a new category that is like the data. All these models have their advantages and characteristics.

The dataset for the implementation of the proposed solution is made based on a dataset taken from [54] and from the department of agronomy Islamia University Bahawalpur (IUB) Pakistan. Each crop and soil have unique nutrient requirements. Furthermore, the fertilizer needs to be used according to the existing soil NPK nutrient level. Therefore, the dataset is used to recommend the appropriate fertilizer according to the soil, type, crop type, and existing soil NPK nutrient level. As the available dataset is small; therefore, the deep learning approaches are not suitable for the small dataset size. The dataset is partitioned into 80:20 ratios in terms of training and testing. The Sciket learn python library is used for the implementation of the machine learning model. The dataset is firstly preprocessed for missing and inappropriate values in the dataset.

For each set of input tuples (X), the appropriate fertilizer (y) is recommended according to the input features. The objective of the machine learning model is to recommend appropriate fertilizer (y) from the set of input conditions 'X', as expressed by Eq. 1, where 'X' is the set of input features.

F(X)→y(1)

View Source 'X' is the set of input features expressed by Eq. 2, where 'St' is the soil type, 'Ct' is the crop type and 'Fl' is the soil fertility level. X={St,Ct,Fl(2)

View Source 'Y' is the set of commercially available fertilizers name used as the output of the model for each tuple of input combinations. Each commercially available fertilizer has different compositions of NPK nutrients usually expressed in their names. The set of fertilizers names (y) used for the implementation of the proposed solution is expressed by Eq. 3, where 'Py' is used for the 'DAP' fertilizer, 'Qy' is used for the '14-35-14' fertilizer, 'Ry' is used for the '26-28' fertilizer, 'Ty' is used for '10-26-26' fertilizer name, 'Uy' is used for the 'Urea' fertilizer, 'Vy' is used for the '20-20' fertilizer and 'Wy' is used for the '17-17-17' fertilizer. The distribution of the selected fertilizers in the dataset is shown in Figure 5.

y={Py,Qy,Ry,Ty,Uy,Vy,Wy(3)

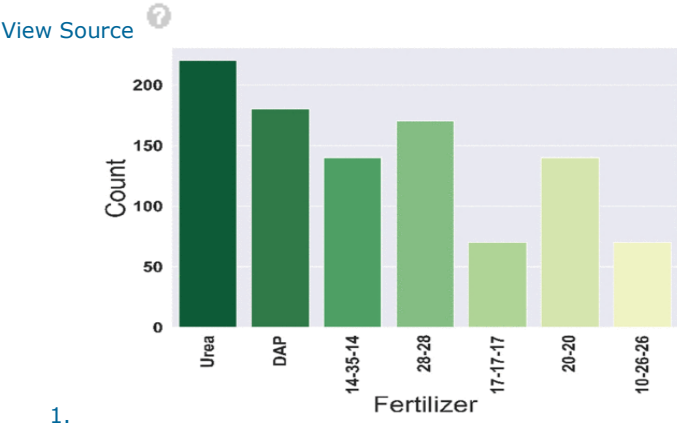


FIGURE 5. Distribution of fertilizer names in the dataset.

Show All

To find the relationship between the input feature and the output (fertilizer name) of the machine learning model, the fertilizer names are classified in Table 1.

TABLE 1 Classification of Soil Nitrogen (N) Level

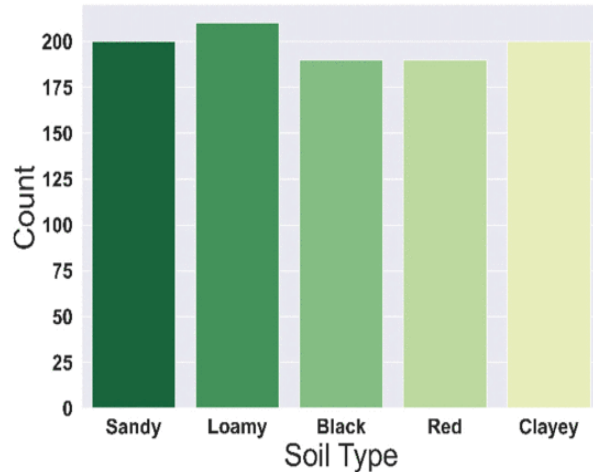
FERTILIZER NAME	CLASS
Urea	1
Di-Ammonium Phosphate (DAP)	2
14-35-14	3
26-28	4
17-17-17	5
20-20	6
10-26-26	7

Each soil has its unique characteristics and composition. Moreover, different types of soils have different capabilities to hold and supply nutrients to plants. The type of soil is an important factor for the fertilizer recommendation. Therefore, the fertilizer

recommendation is also adjusted according to the soil type. For the implementation of the proposed solution, the set of soil types (S) is given in Eq. 4 where Sy is used for sandy soil, Ly is used for loamy soil, Bl is used for black soil, Rd is for red soil, and Cy for clayey soil. The distribution of the soil type (S) in the dataset is shown in Figure 6.

$$St=\{Sy,Ly,By,Rd,Cy\}(4)$$

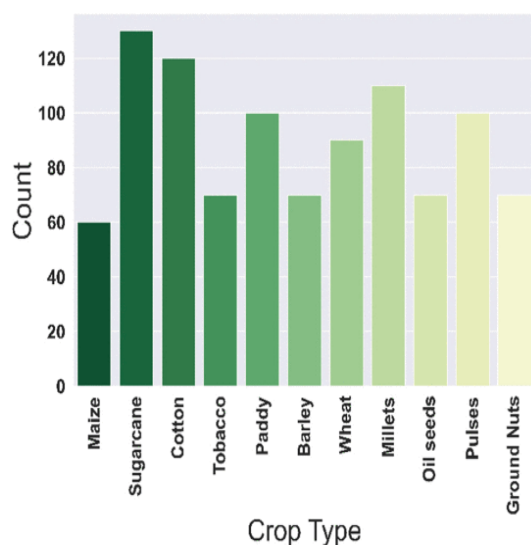
[View Source](#) <sup>?</sup>



**FIGURE 6.**  
Distribution of soil type in the dataset.

[Show All](#)

Each crop has its requirements of nutrients for successful growth and better production. Some crops require low levels of nutrients, while others require more. Therefore, along with the soil type, the crop type is also an important factor for crop fertilizer requirements and recommendations. The distribution of the selected crop in the dataset is shown in Figure 7.



**FIGURE 7.**  
Distribution of crop type in the dataset.

[Show All](#)

The set of crop types (Ct) used for the implementation of the proposed fertilizer recommendation is expressed in Eq. 5, where 'Mz' is used for maize, 'Sc' is used for sugarcane, 'Cn' is used for cotton, 'Tb' is used for tobacco, 'Pd' is used for paddy, 'Br' is used for Barley, 'Wt' is used for wheat, 'Ml' is for millets, 'Os' for oil seeds, 'Pl' is for pulses and 'Gn' for the ground nuts crop.

$$Ct=\{Mz,Sc,Cn,Tb,Pd,Br,Wt,Ml,Os,Pl,Gn\}(5)$$

[View Source](#) <sup>?</sup>

The soil fertility level (Fl) is defined by the level of nitrogen (N), phosphorous (P), and Potassium (K) in soil expressed by Eq. 6.

$$Fl=\{N,P,K\}(6)$$

[View Source](#) <sup>?</sup>

For implementation, the level of three macronutrients (N, P, and K) in soils are classified into 'low', 'medium', and 'high' according to their concentration in the soil. The classification of N, P, and K nutrient levels in the soil is given in Table 2, Table 3, and Table 4, respectively.

**TABLE 2** Classification of Soil Nitrogen (N) Level

NITROGEN (N) LEVEL	CLASS
>10	Low
10-50	Medium
<50	High

**TABLE 3** Classification of Soil Phosphorous (P) Level

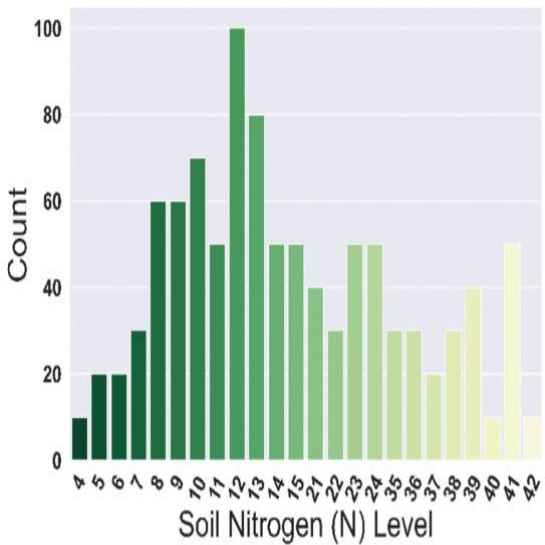
PHOSPHOROUS (P) LEVEL	CLASS
>10	Low
10-100	Medium
<100	High

**TABLE 4** Classification of Soil Potassium (K) Level

NITROGEN (N) LEVEL	CLASS
>10	Low
10-15	Medium
>15	High

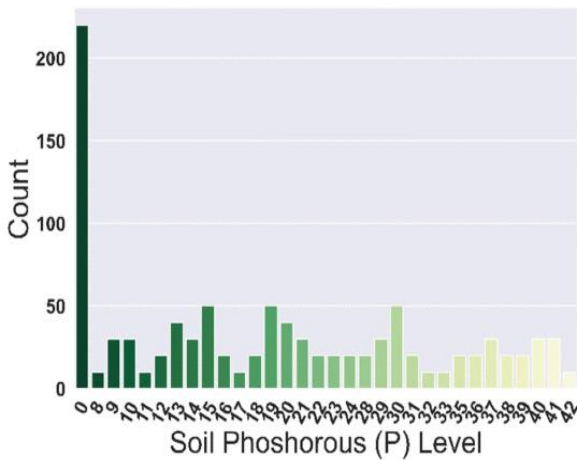
The distribution of the three macronutrients in the dataset is classified according to the described classification. The soil nutrients are observed in milligrams of nutrients per kilogram of soil (mg/kg). The distributions of Nitrogen

(N), Phosphorus (P), and Potassium (K) in the dataset are shown in Figure 8, Figure 9, and Figure 10, respectively.



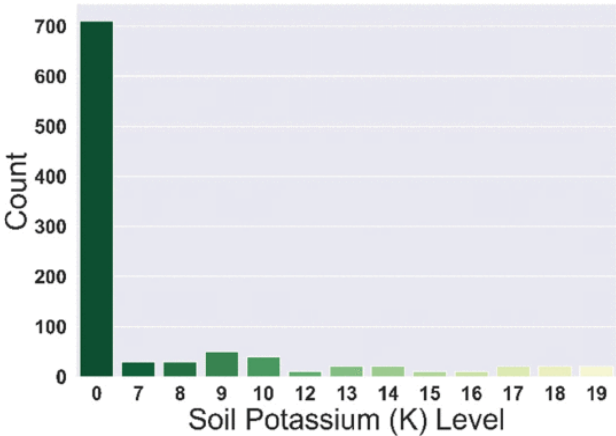
**FIGURE 8.** Distribution of soil Nitrogen (N) level in the dataset.

[Show All](#)



**FIGURE 9.** Distribution of soil Phosphorous (P) level in the dataset.

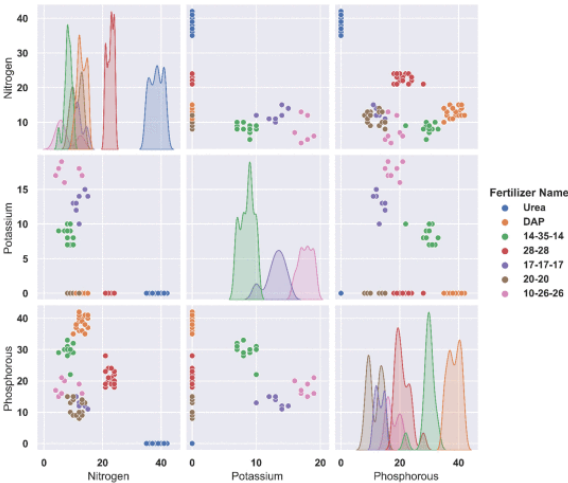
[Show All](#)



**FIGURE 10.** Distribution of soil Potassium (K) level in the dataset.

[Show All](#)

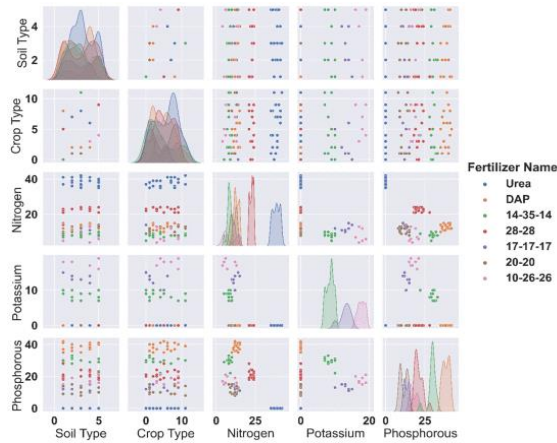
The usage of fertilizer is heavily dependent on the crop type, soil type, and the existing soil fertility in terms of soil nutrients (NPK). The existing soil nutrient level is co-related with crop type, soil type, and fertilizer requirements. The relationships between the existing soil nutrient level with the fertilizer requirements are shown in Figure 11. The relationships between the crop type, soil type, and the existing nutrient level are even more complex. The relationship between all the input features with the fertilizer name is shown in Figure 12.



**FIGURE 11.** Relationship between soil macronutrients and fertilizer.

[Show All](#)





**FIGURE 12.** Relationship between input features and fertilizer requirements.

[Show All](#)

## SECTION IV.

### Results

The proposed solution is evaluated in terms of the accuracies of soil fertility mapping along with a machine learning model to recommend an appropriate amount of fertilizer according to the soil fertility level, crop type, and soil type.

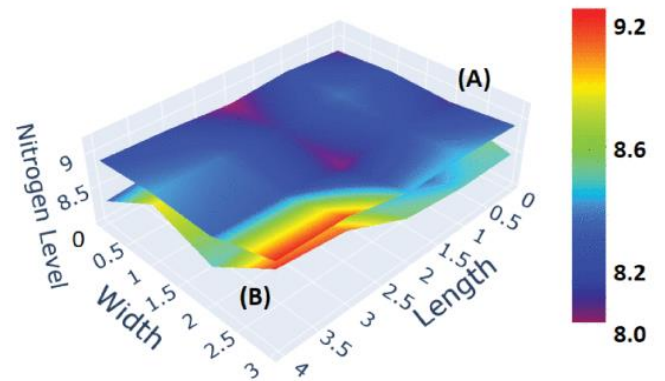
#### A. Accuracy of Soil Fertility Mapping

Initially, the accuracy of the soil fertility mapping by the proposed solution is evaluated. The accuracy of the proposed solution is observed by the accuracy of the soil fertility observations in terms of macro-elements (NPK). The NPK observation by the proposed solution is compared with the soil chemical analysis to determine the accuracy of soil fertility mapping. In terms of comparison, the soil fertility is observed in an experiment area of one acre by the proposed solution along with the standard method to observe the difference at equally distanced twenty sample points.

The NPK observations by the proposed solution and chemical analysis methods in the experiment area are analyzed for each of the three macro elements (NPK) of soil fertility. For comparison purposes, the Bald Man Altman difference plot is used to compare the difference in observations by two different methods or instruments for the same measurement or observation.

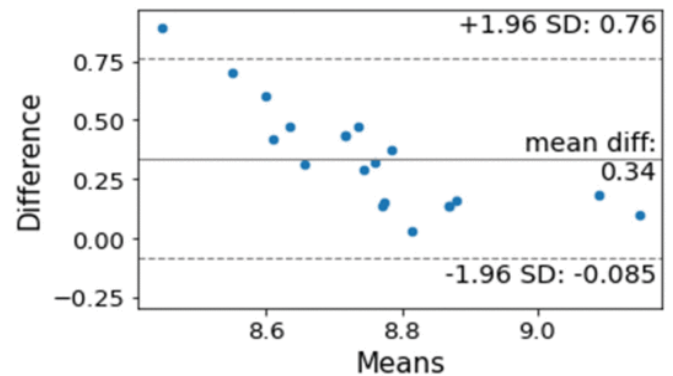
The Nitrogen (N) observations by the proposed solution and the standard method of soil chemical analysis are shown in parts A and part B of Figure 13, respectively. The Nitrogen (N) observations by the proposed solution and the standard chemical-based approach are in line with each other at each of the twenty-sampling points. The mean difference in observations by the two methods, determined by the Bald Altman difference plot, comes out to be 0.34 as shown in Figure 14. Thus, the mean difference between the two methods for each sample point observation is 0.34. This means that the proposed solution in terms of the average measure's nitrogen level is 0.34 more at each of the sample points as compared to the standard soil chemical analysis method. The mean difference of 0.34 in Nitrogen(N) observations by both methods reveal that the proposed solution is accurate in

soil nitrogen (N) observation. The Kernel Density Estimate (KDE) of Nitrogen (N) observation by two both methods are shown in Figure 15, which also reveals the similarity of Nitrogen (N) observation by both methods.



**FIGURE 13.** Distribution of Nitrogen (N) level in experiment area A. by proposed method B). by standard soil chemical analysis.

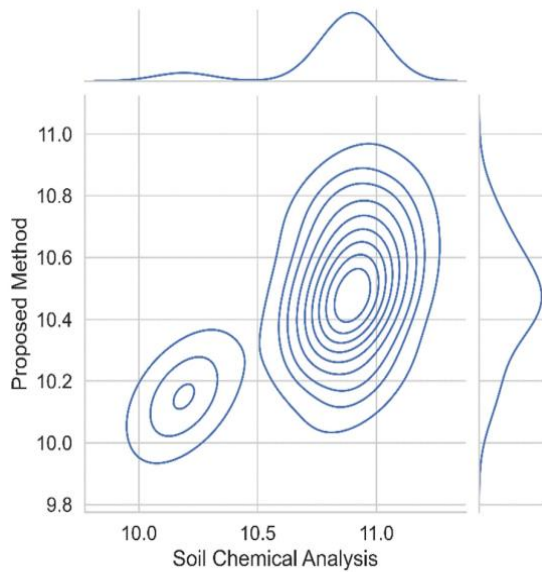
[Show All](#)



**FIGURE 14.** Bland-Altman Plot for the mean difference in Nitrogen (N).

[Show All](#)

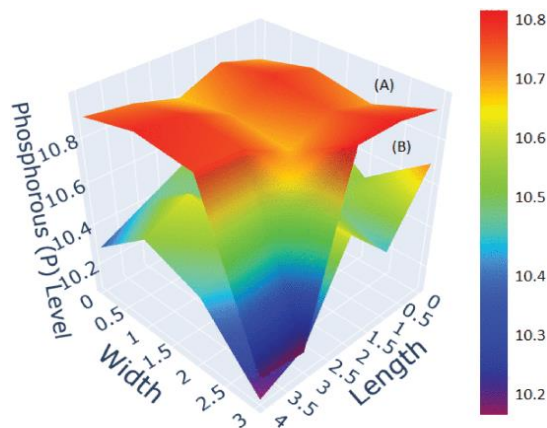




**FIGURE 15.**  
KDE plot of Nitrogen observations (N) by two methods.

[Show All](#)

The Phosphorous (P) observations by the proposed solution and the standard method of soil chemical analysis are shown in parts A and part B of Figure 16, respectively.

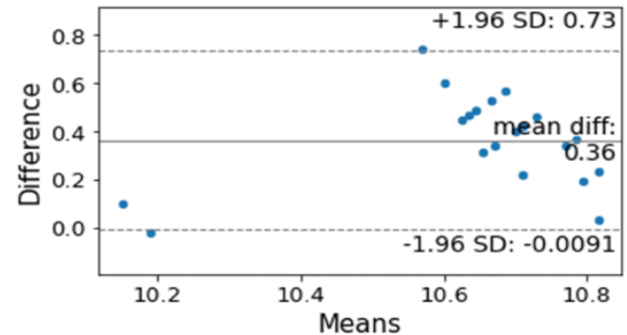


**FIGURE 16.**  
Distribution of Phosphorous (P) level in experiment area A. by proposed method B). by standard soil chemical analysis.

[Show All](#)

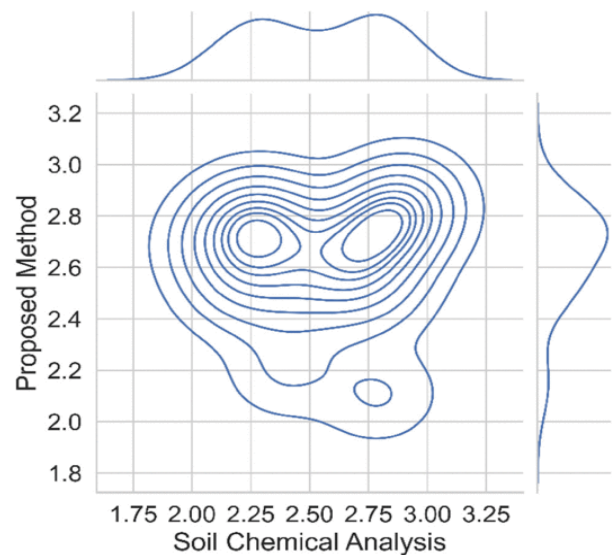
The Phosphorous (P) observations mapping by the proposed solution and the standard chemical-based approach reveal that both observations are in line with each other at each sampling point. The mean difference in observations by both methods, determined by the Bald Altman difference plot, is 0.36, as shown in Figure 17. The mean difference between both methods for each sample point observation is 0.36. This means the proposed solution in terms of the average measure's Phosphorous (P) concentration is 0.36 more at each of the sample points as compared to the standard soil chemical analysis method. The mean difference of 0.34 in Phosphorous (P)

observations by both methods reveals that the proposed solution is accurate in soil nitrogen observation. The Kernel Density Estimate (KDE) of Phosphorous (P) observation by two both methods are shown in Figure 18, which reveals the similarity of Phosphorous (P) observation by both methods.



**FIGURE 17.**  
Bland-Altman Plot for mean difference in Phosphorous (P) observations.

[Show All](#)

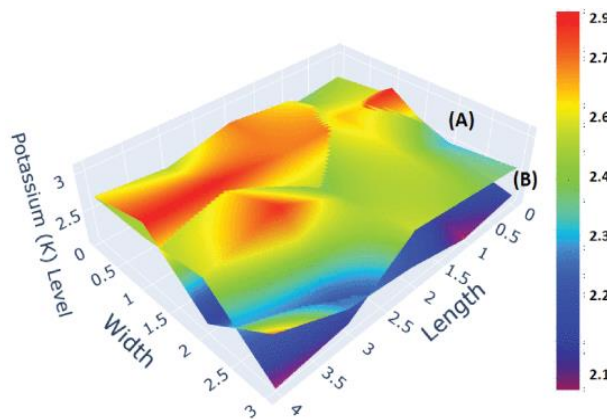


**FIGURE 18.**  
Kernel density estimate (KDE) plot of Phosphorous (P) observations by two methods.

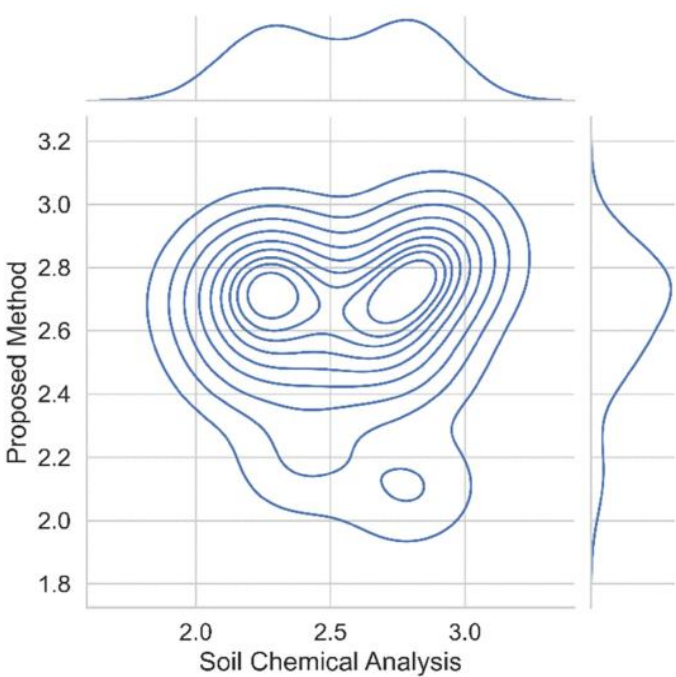
[Show All](#)

The Potassium (K) observations by the proposed solution and the standard method of soil chemical analysis are shown in parts A and part B of Figure 19, respectively. The Potassium (K) observations mapping by the proposed solution and the standard chemical-based approach reveal that both observations are in line with each other at each sampling point. The mean difference in observations by both methods, determined by the Bald Altman difference plot, is 0.36, as shown in Figure 20. The mean difference between the two methods for each sample point observation is -0.13. This means the proposed solution in terms of average measured Potassium (K) concentration is -0.13 at each of the sample points as compared to the standard soil chemical analysis method. The mean

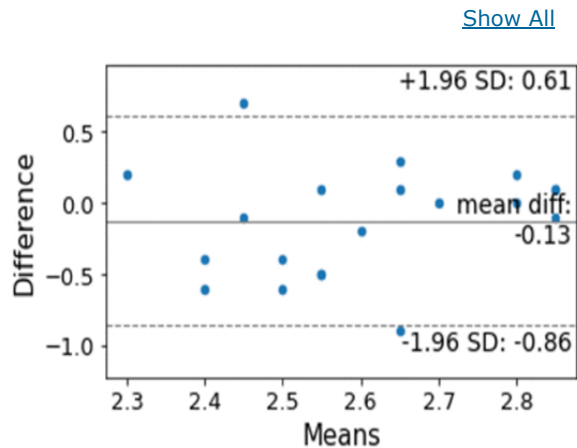
difference of 0.13 in Potassium (K) observations by both methods reveals that the proposed solution is accurate in soil Potassium (K) observation. The Kernel Density Estimate (KDE) of Potassium (K) observation by both methods is shown in Figure 21, which reveals the similarity of Phosphorous (P) observation by both methods.



**FIGURE 19.** Potassium (K) observation A) By Proposed solution B) By chemical analysis.



**FIGURE 21.** Kernel density estimate (KDE) plot of Potassium (K) observations by two methods.



**FIGURE 20.** Bland-Altman Plot for mean difference in Potassium (K) observations.

**B. Performance of the Machine Learning Model**

The machine learning model is used to recommend the fertilizer according to the soil fertility level, crop type, and soil type. The machine learning models are evaluated based on accuracy, precision, and recall. Moreover, the accuracies of the three machine learning models are also evaluated.

**1) Accuracy of the Machine Learning Model**

The accuracy of the model is the number of correct predictions out of the total prediction made by the machine learning model, as expressed in Eq. 7. The accuracy of each evaluated model from training and test datasets is given in Table 5. It is observed that GNB perform better in the case of the accuracy of appropriate fertilizer recommendation according to the context with 97% and 96% accuracy from training and test dataset respectively.

$$\text{Accuracy(\%)} = \frac{\text{Correctre recommendation}}{\text{Total recomendation}} \times 100(7)$$

[View Source](#)

**TABLE 5** Accuracy Comparison of Models

Model	Training Accuracy (%)	Testing Accuracy (%)
LR	94	92
GNB	97	96
SVM	96	94
KNN	95	93

## 2) Confusion Matrices

Precision is the ratio of True positive (Tp) to the sum of the 'Tp' and false positive (Fp), expressed by [Eq. 8](#). Recall is the ratio of the 'Tp' to the sum of 'Tp' and False Negative (Fn) expressed by the [Eq. 9](#). F1 score is expressed by [Eq. 10](#).

$$\text{Precision} = \text{Recall} = \text{F1} = \frac{\text{Tp}}{\text{Tp} + \text{Fp}} = \frac{\text{Tp}}{\text{Tp} + \text{Fn}} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)(9)(10)$$

[View Source](#) 

The fertilizer names are encoded into numerical labels to facilitate the reporting of confusion matrices and classification reports, as shown in Table 6.

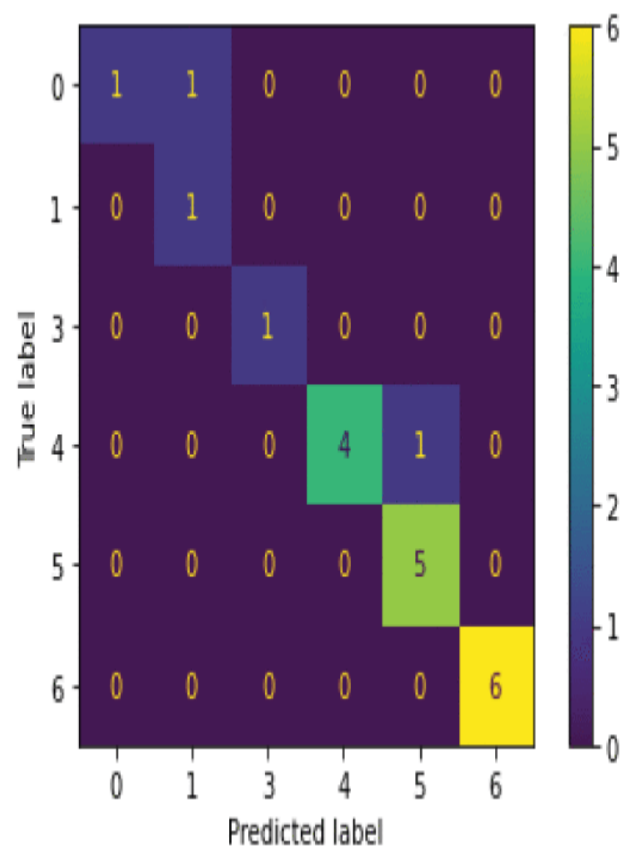
**TABLE 6** Numerical Encoded Labels of Soil Nitrogen (N) Level

FERTILIZER NAME	LABELS
Urea	0
Di-Ammonium Phosphate (DAP)	1
14-35-14	2
26-28	3
17-17-17	4
20-20	5
10-26-26	6

The confusion matrix of the LR model for the fertilizer prediction is shown in Figure 22, and the classification report is given in Table 7.

**TABLE 7** Classification Report of Logistic Regression (LR) Model

Fertilizer Name	Label	Precision	Recall	F1-Score
Urea	0	1.0	0.50	0.67
DAP	1	0.50	1.0	0.67
14-35-14	2	1.0	1.0	1.0
26-28	3	1.0	0.80	0.89
17-17-17	4	0.83	1.0	0.91
20-20	5	1.0	1.0	1.0
10-26-26	6	1.0	1.0	1.0



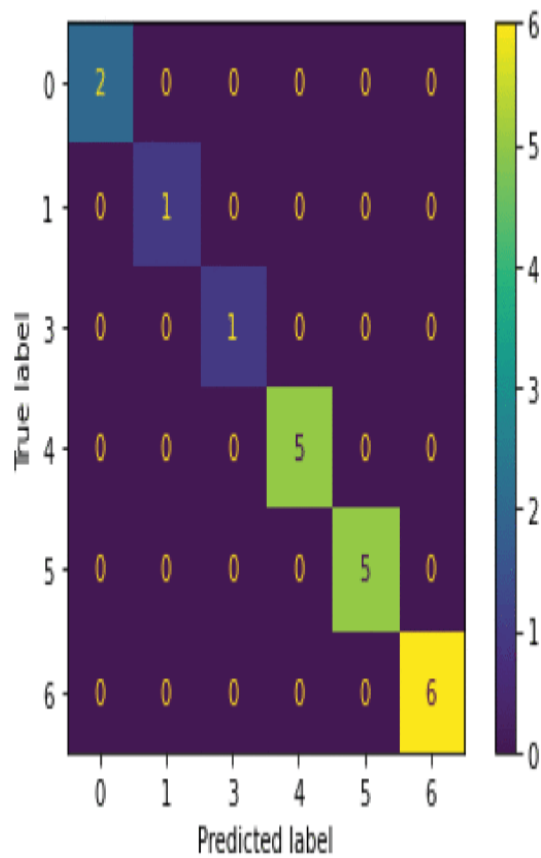
**FIGURE 22.** Confusion Matrix of logistic regression.

[Show All](#)

The confusion matrix of the GNB model for the fertilizer prediction is shown in Figure 23, and the classification report is given in Table 8.

**TABLE 8** Classification Report of GNB

Fertilizer Name	Label	Precision	Recall	F1-Score
Urea	0	1.0	0.7	0.76
DAP	1	1.0	0.9	1.0
14-35-14	2	0.95	1.0	1.0
26-28	3	1.0	1.0	1.0
17-17-17	4	0.83	0.8	0.9
20-20	5	1.0	0.9	0.87
10-26-26	6	1.0	1.0	0.9



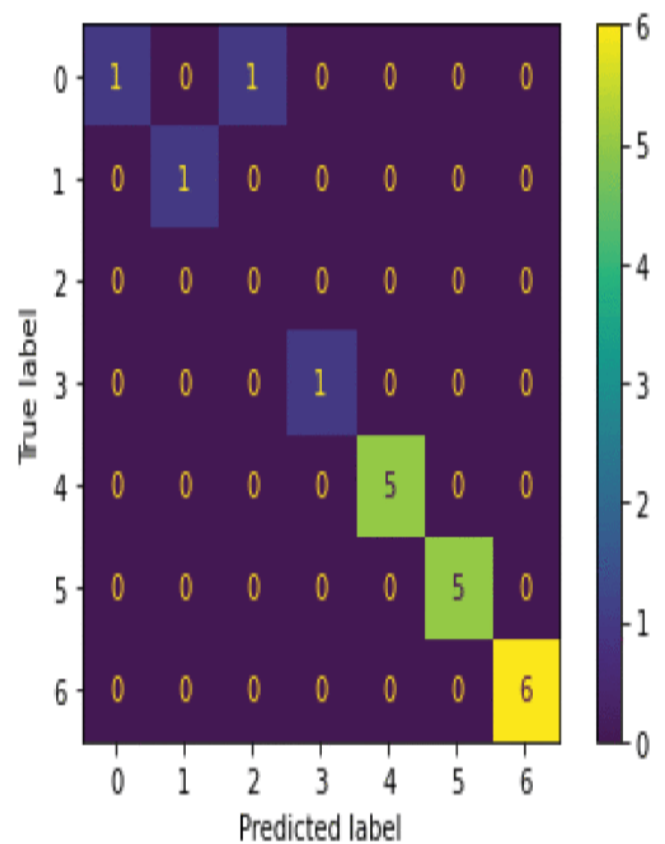
**FIGURE 23.**  
Confusion Matrix of Gaussian Naïve Bayes (GNB).

[Show All](#)

The confusion matrix of the SVM model for the fertilizer prediction is shown in Figure 24, and the classification report is given in Table 9.

**TABLE 9** Classification Report of SVM Model

Fertilizer Name	Label	Precision	Recall	F1-Score
Urea	0	1.0	0.80	0.69
DAP	1	0.70	1.0	0.77
14-35-14	2	1.0	0.9	1.0
26-28	3	0.9	0.80	0.85
17-17-17	4	0.93	1.0	0.85
20-20	5	1.0	0.9	1.0
10-26-26	6	1.0	0.8	0.9



**FIGURE 24.**  
Confusion Matrix of Support Vector Machine (SVM) model.

[Show All](#)

The confusion matrix of the KNN model for the fertilizer prediction is shown in Figure 25, and the classification report is given in Table 10.

**TABLE 10** Classification Report of the K-Nearest Model (KNN) Model

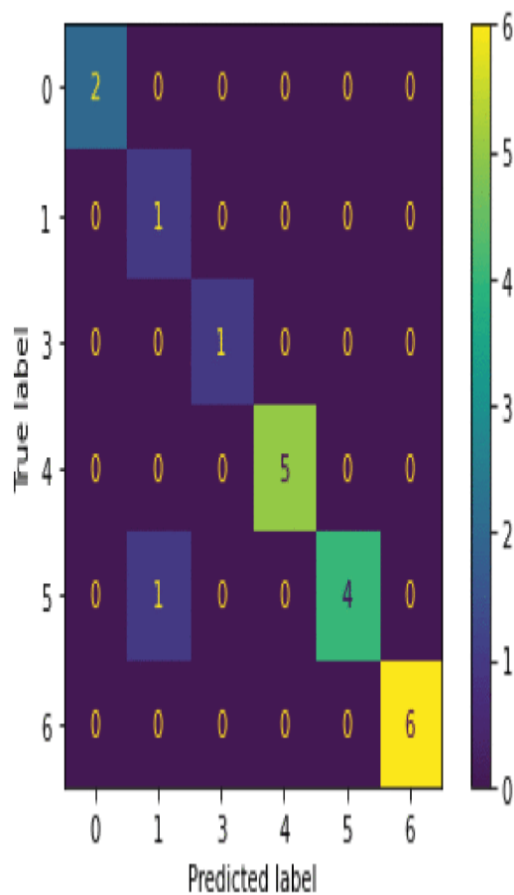


Fertilizer Name	Label	Precision	Recall	F1-Score
Urea	0	0.9	1.0	0.77
DAP	1	0.80	0.9	0.87
14-35-14	2	0.8	1.0	0.97
26-28	3	1.0	0.90	0.98
17-17-17	4	0.93	1.0	1.0
20-20	5	0.9	0.89	0.87
10-26-26	6	0.81	0.92	1.0

fertility context captured through IoT-assisted soil fertility mapping. The soil fertility observations by the proposed soil fertility mapping were in line with standard soil chemical analysis with the mean difference of 0.34, 0.36, and -0.13 for Nitrogen (N), Phosphorous (P), and Potassium (K) observations, respectively. The evaluation of the proposed model for fertilizer recommendation reveals that GNB was the most accurate machine learning model with accuracies of 96% and 94% for training and test datasets. Due to the availability of limited datasets, the deep neural network was not applicable. Therefore, the addition of a new dataset with a deep learning application can be a good contribution in the future.

#### ACKNOWLEDGMENT

The authors are also grateful to the Department of Computer Science, College of Computing, Khon Kaen University, Thailand, for supporting this study.



**FIGURE 25.**  
Confusion Matrix of K-Nearest Neighbor (KNN).

[Show All](#)

The results of confusion matrices and classification reports reveal that the GNB model is better in terms of precision, recall, and F1 score for fertilizer predictions.

#### SECTION V.

##### Conclusion

A machine-learning-based fertilizer recommendation system is proposed. The appropriate fertilizer is proposed according to the soil type, crop type, and real-time soil