iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** VIII **Month of publication:** Aug 2023

DOI: https://doi.org/10.22214/ijraset.2023.55225

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue VIII Aug 2023- Available at www.ijraset.com

Soil Fertility Prediction

Aishwarya Patil[1], Varshini Kulkarni[2], Sachin Desai[3]
Department of Master of Computer Applications, VTU/KLS Gogte Institute of Technology, Belagavi

Abstract: ***Soil fertility plays a important role in figuring out agricultural productiveness and sustainability. Traditional methods of assessing soil fertility involve time-consuming and high-priced laboratory tests, restricting their scalability and real-time applicability. To overcome these challenges, this study proposes a data-driven method utilizing machine learning techniques for accurate and efficient soil fertility prediction. Several machine learning algorithms, inclusive of decision trees, Random Forests, k nearest neighbors, and Gradient Boosting Machines (GBM), are employed to model the complex relationships among soil properties and fertility. Feature selection techniques are carried out to identify the most influential soil parameters for enhanced prediction accuracy and reduced model complexity.The outcomes demonstrate that machine learning models can appropriately predict soil fertility, outperforming traditional approaches in terms of speed and cost-effectiveness. Moreover, the characteristic selection process identifies key soil properties that have the most significant effect on fertility, offering valuable insights for agricultural decision-making and targeted soil management. The proposed approach offers potential applications in precision agriculture, enabling farmers to make knowledgeable choices regarding crop selection, nutrient management, and irrigation strategies based totally on actual-time soil fertility predictions. By optimizing resource allocation and minimizing environmental influences, this data-driven solution contributes to the promotion sustainable agricultural practices and guarantees food safety for a growing global population. These are the essential nutrients that the crop requires for its growth pH nitrogen(N), phosphorus(P), potassium(K), CaCo3, Organic Carbon, Organic matter, CEC (Cation exchange capacity) .***

Keywords: Soil, Prediction, Fertile, Non Fertile

I. INTRODUCTION

Agriculture is one of the predominant career in India. huge populace of India relies upon agriculture as their most important source of profits. With time, the demand for manufacturing has been multiplied exponentially[2].

one of the enormous elements that has an instantaneous impact on crop production and first-class is soil fertility and its nutrient control. offering vegetation with the suitable amount

of nutrients on the right time is the key to a success crop production.[1]

Any malfunction of the soil homes impacts now not simplest agriculture however also the water cycle of the earth and to an extent paves way for herbal calamities. The parameters that have an effect on the fertility are many, that are predicted in laboratories using traditional methods, and its want of the

hour to revolutionize those estimation strategies of soil fertility using automated strategies.[5]

To have an powerful production of the crop and adding fertilizers inside the proper ratio within the soil, it is important for the farmers to realize the soil nutrient composition. consequently, strength of soil nutrient evaluation has grow to be the need of this today's world.The machine learning algorithm plays a high function and offers faster and correct consequences[6].

Machine learning algorithms like Random forest, k-nearest neighbour, gradient boosting were used for results in soil nutrient analysis.

The motto is to expect the outcomes accurately thru the implementation of the ML approach.[6]

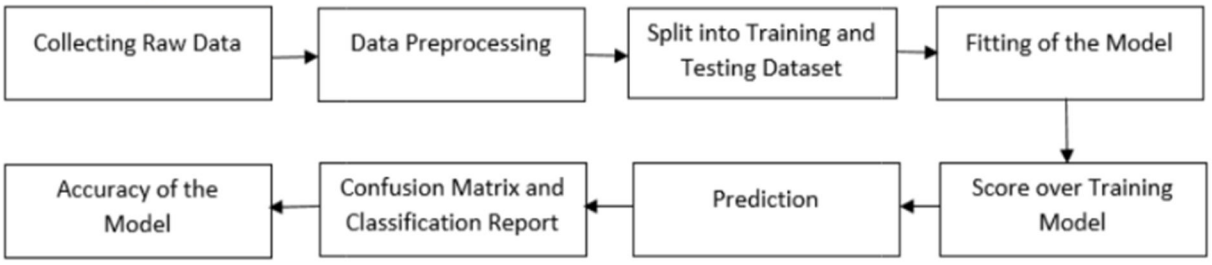


Fig 1. Methodology

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue VIII Aug 2023- Available at www.ijraset.com

II. PROPOSED SYSTEM

The proposed system predicts the amount of soil nutrients specifically N,P,K the use of the trained data set which has various parameters and the more than one linear regression set of rules.for this reason, understanding those values of soil nutrients, it might be easier for the farmers for adding fertilizer in a proper variety and bring higher yield.[6]

1) *Methodology:* Soil fertility is measured in terms of presence/ absence of nutrients present in the soil. chemical properties plays a significant role in deciding the fertility status.The data set is being imported, it consists of parameters such as pH nitrogen(N), phosphorus(P), potassium(K), CaCo3, Organic Carbon, Organic matter, CEC (Cation exchange capacity)those parameters play a massive position in development and increase of a plant. thus, these parameters are used as variables where the nitrogen, potassium and phosphorous bureaucracy the goal variable respectively.[6] The data set is skilled the use of the MLR model.This version is used due to its performance to handle a couple of variables. The scatter plot and pairs characteristic that is used for the advent of graph of the model that shows the contribution of each parameter. There are extraordinary values of each nitrogen, potassium and phosphorous values predicted after giving the parameters pH, moisture, electric conductivity, temperature as an input. The expected values are then used for growing graphs for N, P, K respectively.[6]
2) *Data Pre-processing:* The data set acquired was in row form and consisted of redundant data. some data values were missing. For correct prediction, the rows containing lacking values and the outliers within the data-set have been removed[2
3) *Python Implementation:* After pre-processing of the information, the machine learning algorithms- K nearest neighbour, Random forest and gradient boosting had been implemented on the data set the usage of Python. The classes- good, average and poor have been transformed to numerical form for the ease of prediction. in the k nearest neighbour algorithm prediction has been done.[2]
4) *Accuracy:* The data set become divided into 5 folds and the average accuracy was calculated for each fold. The final accuracy turned into obtained by way of calculating the Average accuracy of the five folds.[2]
5) *Data Visualization:* Data visualization changed was done using the Python library Matplotlib. The variables were additionally converted to numerical form as in line with the requirement.

:

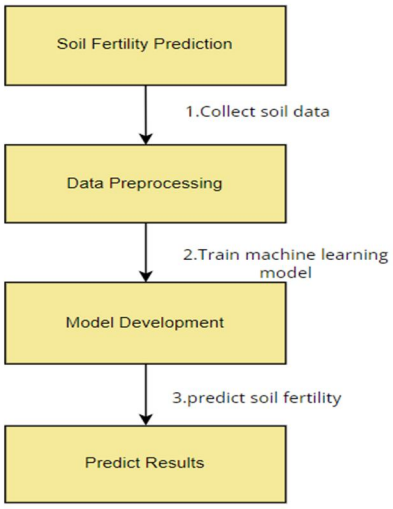


Fig 2data flow diagram

III. ALGORITHM

A. *Random Forest(RM)*

Random forest is a popular algorithm for soil fertility prediction using machine learning. it is a supervised learnin g algorithm that may be used for both type and regression tasks.each decision tree inside the Random forest is grown with the aid of recursively partitioning the data based on different features and splitting criteria. The splits are decided via maximizing the records advantage, Gini impurity, or every other appropriate metric. The tree continues to develop until a distinct termination situation is met.

Fluorescence can be utilized as a valuable tool for assessing soil fertility and health. Specifically, fluorescence spectroscopy is an analytical technique that measures the fluorescence emitted by organic matter present in soil samples when they are exposed to ultraviolet or visible light. the study's finding that induced fluorescence can be used to predict the nitrogen rate directly with an overall accuracy of 0.78 (or 78%) is highly significant. The ability to predict the nitrogen rate directly from soil samples using fluorescence measurements can have practical implications for farmers and agriculture in general[7].

```
----- RANDOM FOREST -----
[1 1 1 1 1 2 0 0 2 2 2 2]
          output
0         Fertile
1         Fertile
2         Fertile
3         Fertile
4         Fertile
5    Medium Fertile
6        Non Fertile
7        Non Fertile
8    Medium Fertile
9    Medium Fertile
10   Medium Fertile
11   Medium Fertile
```

Fig 2: output of random forest

In our test data there are 11 data are present, after applying random forest algorithm, we found 5 are Fertile, 2 are Non Fertile, 5 are Medium Fertile.

B. K Nearest Neighbor

For each instance within the training records, the algorithm calculates the distance between the input feature of the instance and the input features of the data point for which a prediction is required. The most normally used distance metric is Euclidean distance.

The algorithm selects the k nearest neighbors from the training data based on the calculated distances. The value of k nearest neighbors from the training data based on the calculated distances . It determines the variety of neighbors so that it will make a contribution to the prediction.

```
[[5 0 0]
 [0 2 0]
 [0 1 2]]
              precision    recall  f1-score   support

     Fertile       1.00      1.00      1.00         5
  Non Fertile      0.67      1.00      0.80         2
medium Fertile     1.00      0.67      0.80         3

    accuracy                           0.90        10
   macro avg       0.89      0.89      0.87        10
weighted avg       0.93      0.90      0.90        10
```

Fig 3. output of k-nearest neighbour

C. Gradient Boosting

Gradient Boosting is an ensemble device studying approach used for both regression and classification tasks. It builds a strong predictive version with the aid of combining the predictions of more than one vulnerable newbies (usually decision trees) in a sequential way. The idea behind gradient boosting is to accurate the mistakes made with the aid of the preceding model new release via fitting next fashions to the residuals of the preceding ones.

The predictions from the new model are brought to the predictions from the previous version, updating the overall predictions of the ensemble.

```
Accuracy:  1.0
Confusion Matrix:
[[1 0]
 [0 2]]
```

Fig 4: output of gradient boosting

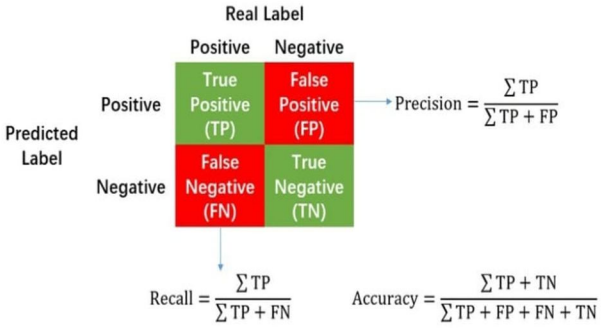


Fig 5:confusion matrix

IV. BENEFITS OF SOIL FERTILITY PREDICTION

A. Crop Yield Optimization

By appropriately predicting soil fertility, farmers can optimize their crop yield capability. Soil fertility prediction fashions can help decide the top of the line quantity and sort of fertilizers wanted for particular crops, making sure that the soil has the important nutrients for surest growth and productiveness. This records enables farmers to make informed choices concerning fertilization practices, leading to accelerated crop yields and advanced aid control.

B. Nutrient Management

Soil fertility prediction can resource in green nutrient management. via predicting the nutrient content and availability inside the soil, farmers can tailor their fertilizer programs to healthy the particular nutrient necessities of vegetation. This reduces the chance of over-fertilization, that can damage the surroundings, whilst making sure that the vegetation get hold of k nutrition for wholesome growth.

C. Environmental Impact Assessment

Soil fertility prediction models can assist check the potential environmental affects of agricultural practices. through thinking about factors inclusive of nutrient runoff and leaching, those models can estimate the risk of nutrient pollutants in nearby water bodies. This records lets in policymakers and land managers to broaden suitable mitigation techniques to minimize environmental harm.

D. Precision Agriculture

Soil fertility prediction may be integrated with different technologies, including far off sensing and geographic records syst ems (GIS), to enforce precision agriculture practices. through mapping soil fertility versions across a area, farmers can follow site-precise fertilizers, making sure that vitamins are dispensed according to the precise needs of every soil vicinity. This targeted ap proach can optimize useful resource usage, reduce fertilizer wastage, and enhance standard crop health and productivity .

E. Land Use Planning and Decision Making

Soil fertility prediction models play a essential role in land use making plans and selection making. through assessing soil fertility characteristics, land managers could make informed selections approximately land suitability for numerous functions, along with agriculture, forestry, or city improvement. Soil fertility prediction allows identify regions with low fertility that may require soil amendments or alternative land uses, stopping capability crop screw ups and maximizing land productiveness.

F. Soil Conservation

Predicting soil fertility is critical for soil conservation efforts. with the aid of figuring out areas with low fertility, erosion-inclined soils, or nutrient imbalances, land managers can implement suitable soil conservation measures. these may additionally consist of terracing, contour plowing, cover cropping, or targeted soil amendments to enhance soil structure, lessen ero sion, and décorate nutrient content.

V. CHALLENGES IN SOIL FERTILITY PREDICTION

1) *Complexity of Soil Systems:* Soil fertility is stimulated by using various factors inclusive of organic depend content material, nutrient levels, soil structure, pH, microbial activity, and greater. The interactions between these elements Make it a complicated system to model appropriately.
2) *Spatial Variability:* Soil fertility can vary significantly over small distances, leading to spatial heterogeneity. Cor rect prediction requires information at a fine spatial decision, which can be costly and time-consumin g to accumulate.
3) *Limited Data Availability:* Soil data collection is an expensive and time-eating procedure. in lots of regions, soil records may be scarce or old, which makes building correct predictive models difficult.
4) *Data Quality and Consistency:* Soil records is often accumulated the usage of diverse strategies and technology, main to inconsistencies in facts exceptional, codes, and scales. Integrating and standardizing statistics from specific assets can introduce noise and errors in the predictions.
5) *Non-Linear Relationships:* The relationships among soil properties and soil fertility signs are frequently non-linear. Traditi onal linear models might not capt ure these complicated relationships correctly.
6) *Over-fitting and Generalization:* Soil fertility prediction models can be prone to over-fitting, mainly while dealing with constrained data or using complicated algorithms. Ensureng the version generalizes properly to new, unseen records is I mportant for realistic packages.

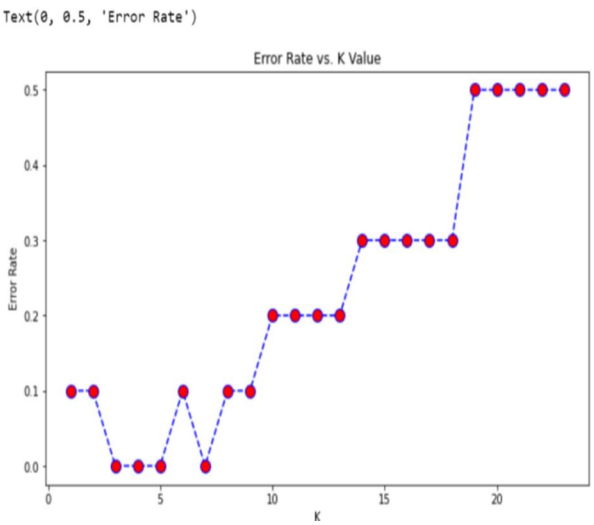


Fig 6: error rate

7) *Temporal Variability:* Soil fertility can change over time because of various natural and anthropological factors, which include weather changes, land use, and agricultural practices. Modeling these dynamic changes accurately is challenging.
8) *Handling Missing Data:* Soil datasets may have missing values, which need to be treated accurately. Managing missing data w ith out introducing biases is a critical aspect of constructing s trong predictive models.

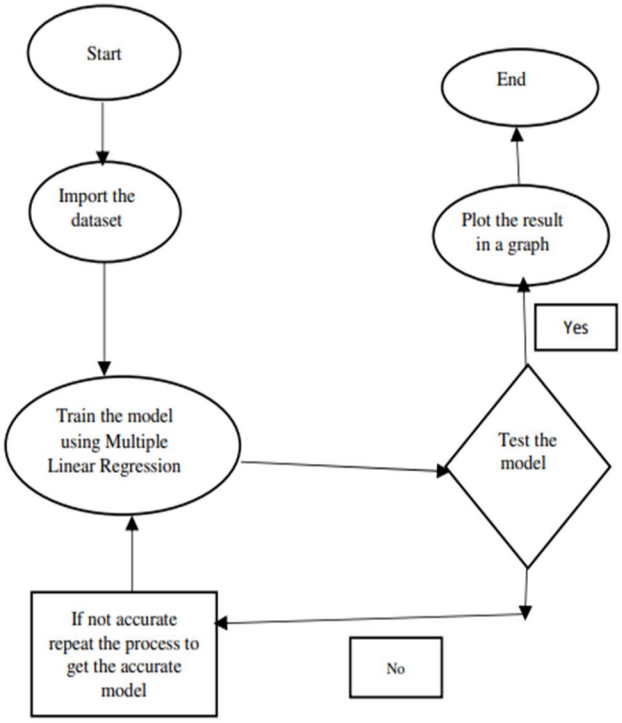


Fig.7 Block Diagram of the Proposed System

VI. FUTURISTIC SOLUTION

Information about pesticides may be Introduced to the crop to improve its growth and to prevent the flowers from being affected by the pests. A graph may be generated that gives the facts about The crops which can be presently planted via the other farmers in a particular area. the use of those facts farmers can get a clean concept about the crop to be planted.[4]

This could further be extended by means of developing a mobile application where the values of N, P, k, CaCo3, organic Carbon, natural matter, CEC that is acquired from the model could be displayed inside the application developed. The values may be stored in a database and displayed in the app. consequently the system advantages the farmers and helps them taking right choices for growing plants.

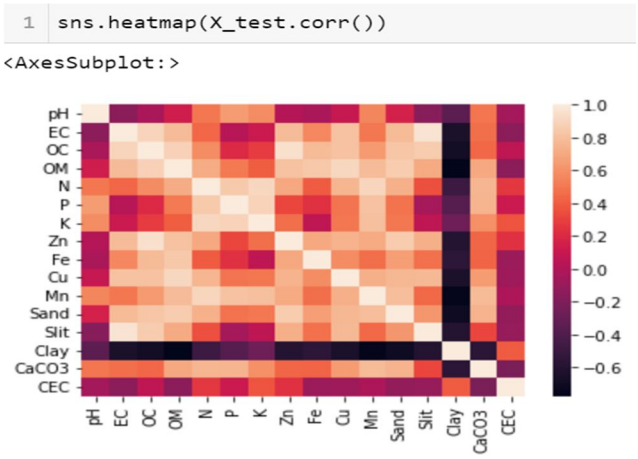


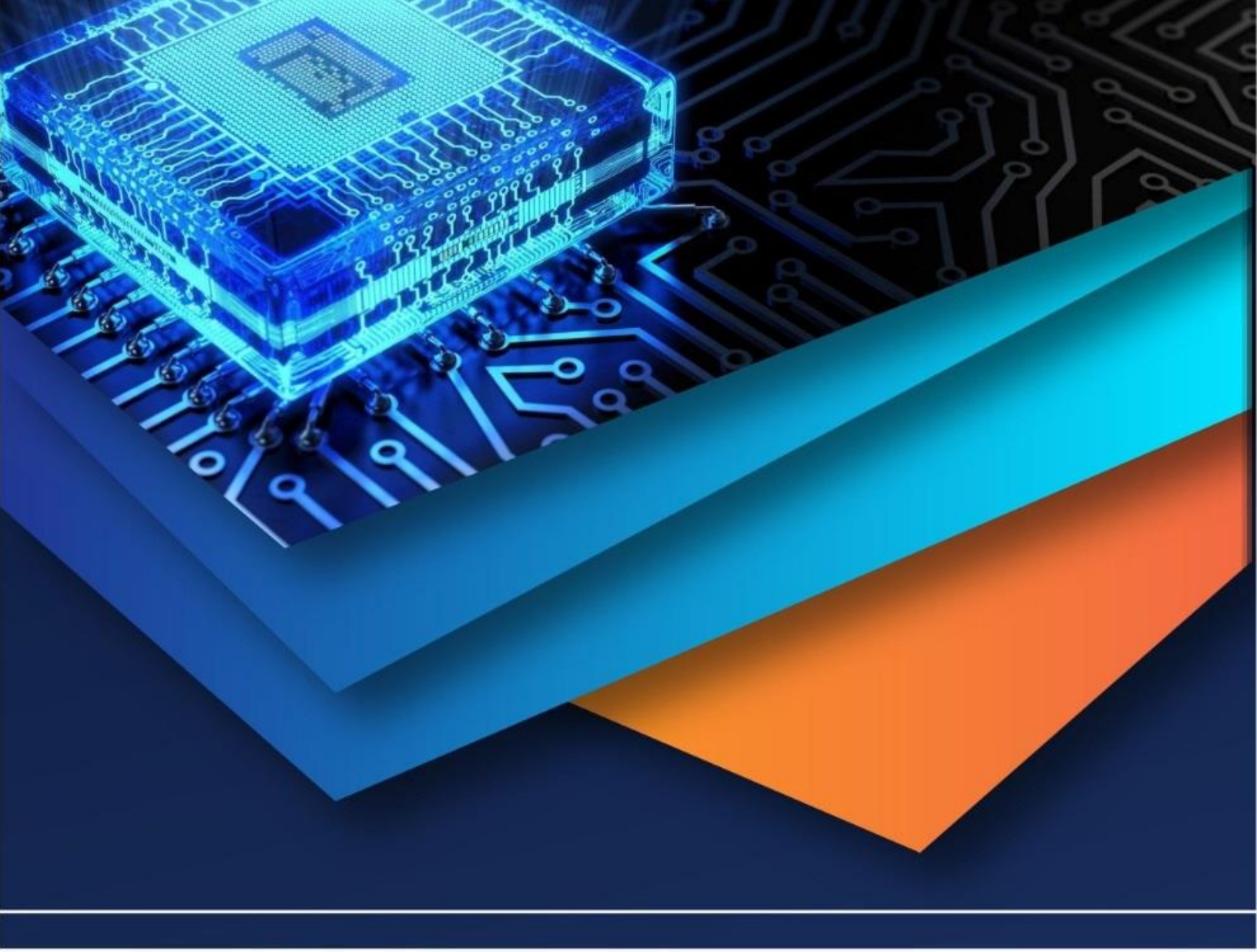
Fig 8: Heatmap

VII. CONCLUSION

soil fertility prediction the usage of machine learning is a promising field that has the potential to improve crop productiveness and yield. through the usage of various machine learning techniques inclusive of k-nearest neighbors, decision trees, Random forest and Gradient boosting, we can analyze soil data which include nutrient content material, pH nitrogen(N), phosphorus(P) and the potassium(k),CaCo3, organic Carbon, organic matter, CEC content present in the soil to appropriately predict soil fertility stag es. This may help farmers make informed selections about fertilization and crop management practices, resulting
in more efficient and sustainable agriculture.Overall, machine learning can be a valuable tool in improving soil fertility management and contributing to sustainable agriculture.

REFERENCES

[1] Jeremy Aliparo, “IOT-BASED ASSESSMENT AND MONITORING OF NPK CONTENT AND FERTILITY CONDITION OF SOIL”, in 2022 IEEE regiin 10 conference, Available: https://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&queryText=soil%20fertility%20prediction
[2] Pranay Malik, “COMPARATIVE ANALYSIS OF SOIL PROPERTIES TO PREDICT FERTILITY AND CROP YIELD USING MACHINE LEARNING ALGORITHMS”, 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) | 978-1-6654-1451-7/20/$31.00 ©2021 IEEE DOI: 10.1109/Confluence51648.2021.9377147, Available:https://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&query Text=soil% 20fertility%20prediction
[3] G Sai Pravallika, “PROFICIENT SMART SOIL BASED IOT SYSTEM FOR CROP PREDICTION”, Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020) Available:https://ieeexplore.ieee.org/search/searchresult.jsp?newsear ch=true&query Text=soil%20fertility%20predictio
[4] P. S. Vijayabaskar, Sreemathi.R, Keertanaa.E, “CROP PREDICTION USING PREDICTIVE ANALYTICS”, 2017 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC) Available:https://ieeexplore.ieee.org/search/sear chresult.jsp?newsearch=true&q ueryTex t=soil%20fertility%20prediction
[5] Sareena Rose , S. Nickolas , S. Sangeetha, “MACHINE LEARNING AND STATISTICAL APPROACHES USED IN ESTIMATING PARAMETERS THAT AFFECT THE SOIL FERTILITY STATUS: A SURVEY”, Second International Conference on Green Computing and Internet of Things (ICGCIoT) Available:https://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&queryText=soil%20fertility%20prediction
[6] Madhumathi R ,Arumuganathan T , Sneha Iyer R , Shruthi R , Shruthhi K,“SOIL NUTRIENT ANALYSIS USING MACHINE LEARNING TECHNIQUES”, In National E-Conference on "Communication, Computation, Control and Automation" (CCCA-2020) Available:https://ieeexplore.ieee.org/search/searchresul t.jsp?newsearch=true&queryText=soil%20fertility%20predictio
[7] Louis Longchamps , Dipankar Mandal, “ASSESSMENT OF SOIL FERTILITY USING INDUCED FLUORESCENCE AND MACHINE LEARNING” ,2022 by the authors. Licensee MDPI, Basel, Switzerland.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY