Journal of Positive School Psychology http://journalppw.com
2022, Vol.6, No.4, 4334-4341

# CROP PREDICTION USING SOIL NUTRIENT BASED ON BEST FIT CROP ANALYSIS [BFCA] ALGORITHM

,

[1]**P.Sakthi Murugan,** Assistant Professor of Computer Science, Kaamadhenu Arts and Science College, Sathyamangalam. Email: sakthigasc@gmail.com
[2]**P.Prabhusundhar,** Assistant Professor of Computer Science, Gobi Arts and Science College, Gobichettipalayam. Email: drprabhusundhar@gascgobi.ac.in
[3]**B.Kiruthika,** Research Scholar, Department of Computer Science, Gobi Arts and Science College. Email: kiruthikabalu@gmail.com
[4]**M.Ramalingam,** Assistant Professor of Computer Science, Gobi Arts and Science College, Gobichettipalayam. Email: ramsgobi@gmail.com

***Abstract*:**

Agriculture is India's economic backbone. Crop cultivation used to be done by farmers with on-the-ground experience. However, crop yields have begun to suffer as a result of climate change. As a result, farmers are unable to select the appropriate crops based on soil and environmental parameters, and the process of manually predicting the appropriate crop/s of land has frequently failed. Crop prediction accuracy leads to higher crop production. Agriculture is entirely dependent on soil fertility, climatic conditions, irrigation, seed quality, harvesting, and other factors.It is really important all around the world. Only a seasoned farmer can recognize the type of soil and select the appropriate crop for it. Predicting the soil type and its surrounding environment for a specific field is crucial for future crop yields. This research focuses on applying a variety of data mining approaches to forecast soil conditions and improve crop yields in the future. The study employs data mining techniques such as clustering, OneR, and J48.These data mining approaches are systematically utilized to anticipate and analyze the bearing of the soil in order to improve crop yield. This technique would be more useful to farmers in identifying the type of the soil and its riches, which, in turn, would help them choose the crop that is best suited to their soil and produces the highest yield.

**Keywords:** Soil Nutrients, Spatial mining, J48, OneR

## 1. INTRODUCTION

Agriculture, which involves the cultivation of crops, is a key contributor to the Indian economy. Crops can be either food or commercial in nature. Paddy, wheat, rice, maize, apple, millets, and other food crops exist alongside commercial crops such as sugarcane cotton, groundnut, and cashew. Weather conditions have a considerable influence on crop productivity [1]. As a result, accurate yield prediction is a significant issue that must be addressed. Farmers would be able to take preventative measures to boost output if yields could be predicted ahead of time. Early forecasting is feasible thanks to the collecting of previous farmer experience, weather conditions, and other influencing elements, which are then stored in a huge database. Rainfall, temperature, humidity, potassium, phosphorus, and nitrogen are all typical input parameters to determine which crop is most suited to the soil. To determine the best crops for profitable yields, create by the Best Fit Crop Prediction Algorithm. For agricultural development, soil is a diversified, nonrenewable, and essential natural resource. It provides minerals, water, and oxygen to plants, which aid in their physical production, vigorous growth, survival, and flourishing.

A strong basis for developing stable and healthy crops is fertile soil [6]. It serves a range of beneficial roles while producing no environmental degradation or harm [7]. The fundamental physical, physiological, biological, and mineralogical features of soil have a direct impact on its fertility [8].As previously stated, chemical testing of manually collected soil samples is typically used to quantify and assess soil attributes. Methods for analyzing or estimating some of the qualities using previously mentioned features are required because the methodology is difficult and time-consuming.

Before the dirt can be portrayed, it should initially be parted into discrete homogeneous classes. Soil examination without a right evaluating is similar to doing handle studies with green plants or lab research with the negligible fundamentals of soil components [9]. Thus, soil arrangement is necessary. Predicting crop yields in front of reap would make it more straightforward for policymakers and ranchers to take fitting advertising and capacity techniques. Beforehand, yield forecast was finished by considering the rancher's information on a specific field and harvest. Ranchers, then again, are compelled to create countless yields as the conditions crumble step by step. Given the present status of issues, a few of them need adequate information on the new harvests and don't seem, by all accounts, to be mindful of the advantages they get from creating them.

This task will help ranchers in deciding the yield of their harvest prior to developing it on the horticultural field, permitting them to settle on the best choices possible.It makes an endeavor to answer the test by building an exceptionally fundamental intuitive forecast framework. It is managed to carry out such a framework with a simple to-utilize interface depending for the most part on web illustrations and furthermore AI rules. This paper depicts a framework as an online application that might utilize information investigation procedures to expect the most worthwhile yield in view of current climate and soil conditions.

## 2. LITERATURE SURVEY

DM approaches have been employed in agriculture by a number of researchers in recent years. The following is a review of how various DM techniques have been used in the field of soil categorization analysis in recent year.

**Murugesan&Radha, [3]** proposed a novelclassification algorithm for effectively classifyingsoil data that combines attribute category rankwith filter-based instance selection. Experimentswere conducted using soil data from the Pollachiarea in Coimbatore district, Tamil Nadu state,India, which is a common marketplace for avariety of grains, vegetables, and fruits. Theproposed model's classification accuracy is alsocompared to that of other classification models.The proposed model has a higher accuracy ratefor soil data, according to the results review. Byselecting the instances, they can define thesignificant attribute category for classifying thesoil data using attribute group rank. Theproposed model has better classificationaccuracy than many other current classifiersunder review, according to the experimentalresearch. In classifying the soil data of thePollachi area, the proposed model has 91.2 percent accuracy, 94.4 percent precision, and94.3 percent recall. The focus of future researchwill be on analyzing the soil types in and aroundCoimbatore. Furthermore, in the future, croppredictionfor specific soil types, as well asweather and climatic conditions, will be needed,which is critical for increasing agriculturalproductivity.

Pandith et al. [5] recommended five supervised machine learning procedures that were applied tothe gathered information: Nave-Bayes, k-Nearest-Neighbor (KNN), Multinomial Logistic Regression, Random-Forest, and Artificial-Neural-Network (ANN). Five models, namely consistency, memory, accuracy, particularity, and-score, were assessed to decide the success of every procedure under survey. Experiments have been directed to decide the most accurate system for foreseeing mustard crop yields. All of the ML techniques under investigation might be utilized to appraise crop yields, as per the discoveries of the experiments. The most elevated precision was predicted by KNN and irregular woods (88.67 %and 94.13 %, individually), while the lowest accuracy was anticipated by Nave Bayes (72.33percent). As far as precision, the most extreme worth expected by ANN was 99.94 percent, while the least worth anticipated by Logistic regression was 24.17%. Yet, for Nave Bayes, all of the classifiers concentrated on expected review upsides of over90%. It says that Nave Bayes had the maximum false negative rate, while Logistic regression had the most reduced genuine negative rate. With specificities of 99.78 percent and 80.72 percent, respectively, and f-scores of 0.9976 and 0.8405,ANN and KNN recorded the most elevated specificity and f-score.

Yield expectation is a vital issue in farming side. Assuming any rancher is intrigued to knowing the amount to yield, he is going to anticipate. Break down the different related properties dataset which like pH, area, and worth from which alkalinity of the not set in stone. Alongside it, level of supplements or dataset like N, P, and K Location is utilized alongside the utilization of outsider applications like APIs for climate and temperature, sort of soil, supplement worth of the

dirt around there, measure of precipitation in the locale, soil. [10] .

Motia& Reddy, [6] proposed the Ensemble Classifier (EC) beat normal classifiers such as DT, KNN, and NB as far as accuracy. For farming soils concentrate on Using a publicly accessible agrarian soil dataset, accuracy ofthree notable characterization models is compared in this review: k-Nearest-Neighbor (k-NN), Naïve-Bayes (NB), and Decision-Tree (DT).Following the examination, an Ensemble Classifier (EC) is proposed, which joins thethree classifiers recently depicted. EC has the highest precision of 84%, rather than k-NN (73.56 %), DT (80.84 %), and NB (72.90 %).As an outcome, it beats the other classifiers. The discoveries propose that EC may be advantageous for precisely characterizing soil types in the agrarian space.

Expectation of Crop Cultivation - R Solanki, D Bein, J Andro-Vasko& W Bein[2].The paper proposes to utilize AI calculations to plan a forecast model for crop yield expectation. On contrasting the presentation of differed direct and non-straight regress or models utilizing 5-crease cross approval, it was observed that utilizing greater part of the fundamental settings, the Random Forest Regress or executed the most straightforward, continued by Nearest-Neighbor Regression, L2 rectilinear relapse with polynomial elements, and backing vector relapse utilizing a Radial Basis Function portion.

Rural information which is being created continually. Thus, horticultural information has which come the enormous information. Smart technologies remember for information assortment electronic gadgets is utilized. In our venture we will break down and mine this horticultural data to get helpful outcomes utilizing innovations like AI and information examination and this outcome will be given to ranchers to get the better crop yield as far as proficiency and efficiency. [11]

Geo area based information collectivity research by yashoyardhan.et.al [13] uncovers that informational indexes applicable to the measurable investigation are gathered from numerous areas. Preprocessing of information incorporates purging and coordination. Purging alludes to the extraction interaction conveyed by extricating important information and taking out the untrustworthy informational indexes. Direction is finished by the strategy applied on the informational indexes and examples. Discernment and evaluation of examples is the approach to portraying the information procured on informational collections advanced [12].

Jay Gholap [14] predicts soil luxury misuse call tree equation. In [15], the maker decided soil characteristics and inspected soil information misuse portrayal methodologies. Soil properties, for example, pH regard, Electrical Conductivity(EC), Potassium, Iron, Copper, etc requested utilizing arrangement estimations [16] use KNN, Naïve Bayes and J48 for separating soil data. This examination recommends the excrement subject to the aspect of supplements found in the soil test set.

## 3. SOIL NUTRIENT ANALYSIS METHODOLOGY

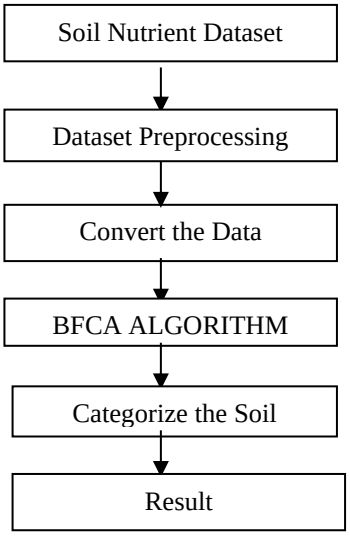
In this system , analyze the soil nutrition parameters are mentioned in the Table 1.

**Table 1: Soil Nutrition data set attributes.**

| ATTRIBUTES | DESCRIPSION |
|---|---|
| N | Nitrogen |
| K | Potassium |
| P | Phosphorus |
| Hum | Humidity of the soil |
| RF | RainFall |
| Tem | Temperature |
| pH | PH value of the Soil |

**Proposed System Work Flow**

The soil nutrientsare passed as an input parameter and dataset has been preprocessed using BFCA algorithm to convert this data fields to predict the crop which is suitable for soil.




For most plants, controlling the pH of the soil is critical to achieving optimal growth circumstances. This is due to the fact that the pH of the soil regulates the solubility of nutrients and harmful metals. As a result, most plants have a preferred pH range in the soil.

When hydrogen ions predominate in a soil, it is said to be acidic. pH, which is defined as the negative logarithm of hydrogen ion activity, is used to express the degree of acidity. As a result, a 0.01-molar hydrogen ion solution's pH is

$$\mathbf{pH = \log -10 \text{ mol } H^{+}/L} \qquad \left( \frac{\quad}{\quad} \right)$$

Table 2 containing the pH values based on the soil nutrient.

**Table 2: pH Range**

| **pH Range** | **Description** |
|---|---|
| Ultra acidic | < 3.5 |
| Extremely acidic | 3.5–4.4 |
| Very strongly acidic | 4.5–5.0 |
| Strongly acidic | 5.1–5.5 |
| Moderately acidic | 5.6–6.0 |
| Slightly acidic | 6.1–6.5 |
| Neutral | 6.6–7.3 |
| Slightly alkaline | 7.4–7.8 |
| Moderately alkaline | 7.9–8.4 |

| | |
|---|---|
| Strongly alkaline | 8.5–9.0 |
| Very strongly alkaline | > 9.0 |

The pH values are described in Table 2 above. If it's less than 3.5, it's extremely acidic. The pH of the soil is neutral if it is between 6.6 and 7.3. Then, if the results are greater than 9.0, the body is highly alkaline.
To determine the amount of phosphorus in the soil [17].

$$\text{Phosphorous(mg/kg)} = \frac{C \times 14}{ODW}$$

Where

C=Phosphorous Concentration
ODW=Oven-dry sample weight(g)
14 =Dilution Factor

## 4. CROP FINDING PROCEDURE

The basic crop-finding process is outlined below. Soil nutrients are an input parameter that should be analyzed to determine which crop is best suited for the soil and should be planted to maximize production..

i. Soil nutrient dataset is imported
ii. The parameters used in the dataset is transformed in particular range is suitable. The redundancy values are reduced and then any of the missing values are removed.
iii. First 80% of the dataset are taken for training samples.
iv. The classification algorithm is used to applied for the training samples.
v. The remaining 20% of the dataset are taken for testing samples.
vi. The trained classifier is used to find the right crop is well suited for the soil. Conclude the crops as a result.

**Best Fit Crop Analysis Algorithm:**

1. Input the soil data set parameters such as N, P, K, Temp, Hum, Rainfall
2. Remove the Noise

   $$\text{Accuracy} = \frac{TP + TN}{TP+TN+FN+FP}$$

   Where TP is the True Positive, TN isTrue Negative,FN is the False Negative,FP is the False Positive.
3. Split the dataset into training and testing
   To find the soil parameter estimationusing the below formula

   $$S_j = \sqrt{\sum_{K=1}^{m} S_k (X_{jk} - X_{ij})^2}$$
4. Retrieve the objectives from deeper layer
5. Collect all the tag from training set.
6. Use the BFCA algorithm to training dataset.
7. Retrieve the characteristics from the training data.
8. Find the test labels
9. Produce the Result that is crops well suited for soil.

## 5.RESULT AND DISCUSSION

The experimental evaluation is done by using dataset. In this dataset Figure1 contains 2200 sample use this samples to predict the crops that is well suited for the soil by using the BFCA algorithm.

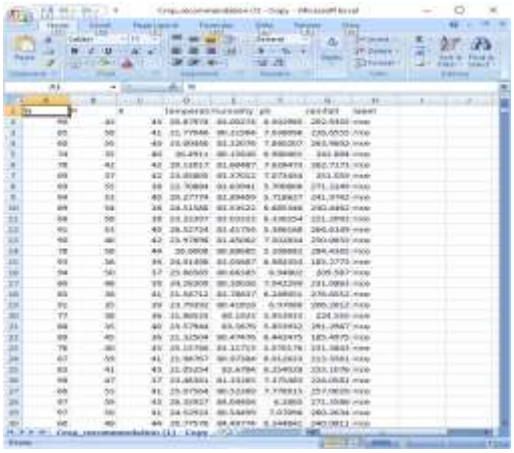


**Figure 1.Soil Nutrient Dataset**

JRIP Rules algorithm uses this dataset to predict the correctly classified instances are 2176 samples and incorrectly classified instance are 24 that are shown in Figure 2.

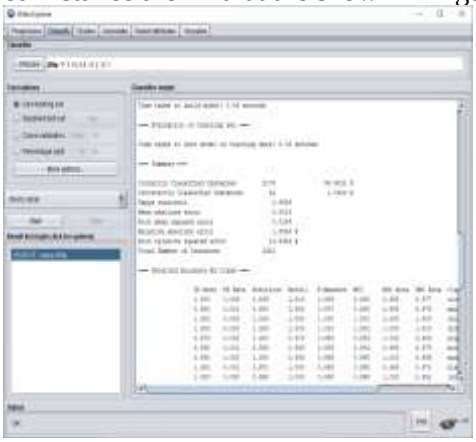


**Figure 2:** OneR Algorithm use the dataset to predict the values contain corrected 1043,In corrected instance 1157

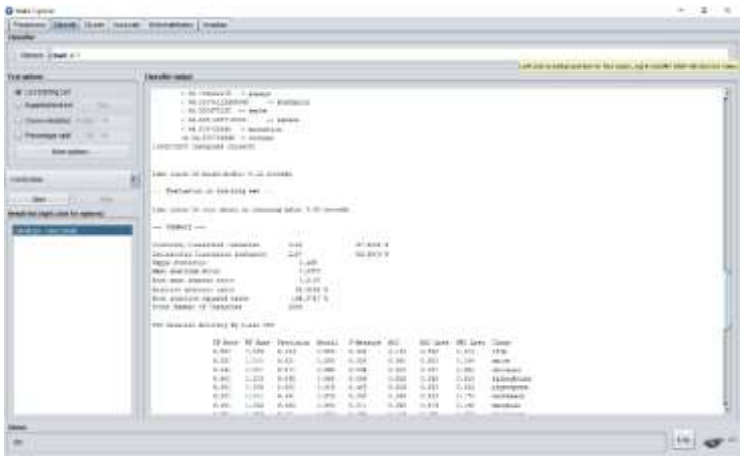


**Figure 3:** Total Instance:2200

**Table 2: Comaprision Between JRip,OneR and BFCA**

| Algorithm | Corrected | In-Corrected | Kappa statistics |
|---|---|---|---|
| JRip | 2176 | 24 | 0.9886 |
| OneR | 1043 | 1157 | 0.445 |
| BFCA | 2198 | 2 | 0.9967 |

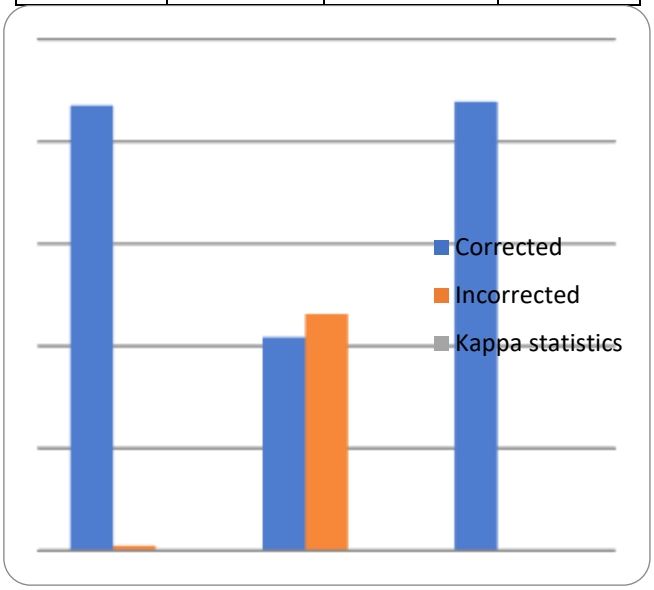



**Figure 4: Result for Classifier Algorithm**

Table 3Compared to the JRip and OneR algorithm, the Best Fit Crop Analysis algorithm is very efficient for the soil nutrient to predict the right crop for the soil. BFCA is the very efficient to predict the crops suitable for the soil nutrients that is shown in Figure 4.

## 6. CONCLUSION

We introduced an algorithm for crop prediction based on soil nutrient characteristics in this research. BFCA algorithm is superior to JRip, OneR, and JRip. With an accuracy of 0.99909 percent, the BFCA is better at predicting soil nutrients and locating crops.

In the future, construct a crop prediction system to determine the best crop for the soil. This method will suggest the best fertilizer based on the soil sample dataset and cropping pattern.

## 7. REFERENCE

1. Food and Agriculture Organization of the United Nations (FAO). World Food Situation: FAO Cereal Supply and Demand Brief. Available online:http://www.fao.org/worldfoodsituation/csdb/en/ (accessed on 17 September 2018).
2. P. S. Nishant, P. SaiVenkat, B. L. Avinash and B. Jabber, "Crop Yield Prediction based on Indian Agriculture using Machine Learning," 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 2020, pp. 1-4, doi: 10.1109/INCET49848.2020.9154036.
3. Murugesan G, Radha DB. Soil dataclassification using attribute group rankwith filter based instance selection model.2020;9(06):7.
4. P.Prabhusundhar, “Distributed Storage Architecture and Hadoop Framework Tools for Crunching Big Data in Clustered Processing System”, TEST Engineering &Management(SCOPUS Journal), ISSN NO. 01913-4120, Vol. 82, Jan-Feb-2020,pp. 4659-4664.
5. Pandith V, Kour H, Singh S, Manhas J,Sharma V. Performance evaluation ofmachine learning techniques for mustardcrop yield prediction from soil analysis.JSR. 2020;64(02):394–398.DOI: 10.37398/JSR.2020.640254.
6. Motia S, Reddy S. Ensemble classifier tosupport decisions on soil classification.IOP Conf. Ser.: Mater. Sci. Eng. 2021;1022:012044.DOI: 10.1088/1757-899X/1022/1/012044.
7. El-Ramady HR, et al. Soil quality and plantnutrition. In Sustainable AgricultureReviews 14, Springer. 2014;345–447.

8. Karlen DL, Ditzler CA, Andrews SS. Soilquality: Why and how?.Geoderma.2003;114(3–4):145–156.
9. Hartemink AE. The use of soilclassification in journal papers between1975 and 2014. Geoderma Regional.2015;5:127–139.
10. Crop Yield Prediction and Efficient use of Fertilizers, S. Bhanumathi ; M. Vineeth ; N. Rohit, 2019 International Conference on Communication and Signal Processing (ICCSP)
11. P.Prabhusundhar, " An Effective Feature Selection and Incremental Classifier Design for Rice Disease Prediction", Journal of Advanced Research in Dynamical & Control Systems (SCOPUS JOURNAL), ISSN NO. 1943-023X, Vol.12, No.4, March-2020.
12. Crop Yield Prediction Using Data Analytics and Hybrid Approach, Shreya V. Bhosale ; Ruchita A. Thombare ; PrasannaG. Dhemey ; Anagha N. Chaudhari, 2018 Fourth International Conference on Computing Communication Control andAutomation (ICCUBEA)
13. Rainfall variability analysis and its impact on crop productivity Indian agriculture research journal 2002 29,33.,8) SPRS Archives XXXVI-8/W48 Workshop proceedings: Remote sensing support to crop yield forecast and area estimates GENERALIZED SOFTWARE TOOLS FOR CROP AREA ESTIMATES AND YIELD FORECAST by Roberto Benedetti A, Remo Catenaro A, Federica Piersimoni B
14. D.Rajesh, ―Application of Spatial Data mining in Agriculture‖ International Journal of Computer Applications, Volume 15, February 2011.
15. Jay Gholap (2012), "Performance Tuning of J48Algorithm for Soil Fertility" ,2012. Asian Journal ofComputer Science and Information Technology 2: 8(2012) 251–252
16. P.Prabhusundhar, " Sensor Data Management and Personal Data Protection in IoT Networking", International Journal of Advanced Science and Technology (SCOPUS Journal), ISSNO NO. 2005-4238 IJAST, Vol.29, No. 7, November – 2020,pp. 105902-105910.
17. Gholap J, Lngole A, Gohil J, Shailesh, Attar V.(2012),"Soil data analysis using classification techniques andsoil attribute prediction", 2012 Jun; 9(3):1–4
18. Suman, Bharat BhushanNaib (2013), "Soil Classificationand Fertilizer Recommendation using WEKA" ,IJCSMSInternational Journal of Computer Science &Management Studies, Vol. 13, Issue 05, July 2013
19. Bray, RU & Kurtz, LT 1945, Determination of total ,Organic and available of phosphorous in soil. Soil Science ,59:39-45.