



International Journal of Marketing Management

ISSN 2454 - 5007



www.ijmm.net

Email ID: editor@ijmm.net , ijmm.editor9@gmail.com

Applying Machine Learning to Soil Classification, Analysis, and Crop Suitability Recommendations in Agriculture

¹ Mr.Khaja Pasha Shaik, ² Shaik Ahsan Ruman, ³ Khaja Yasser Uddin, ⁴ Furqan Ahmed Khan Shaik,

¹ Assistant Professor, Department of CSE-AIML, Lords Institute of Engineering & Technology.

^{2,3,4} Student Department of CSE-AIML, Lords Institute of Engineering & Technology.

Abstract—

Raising agricultural yields to meet the demands of a growing population is a top priority. A number of variables, including soil quality, rainfall, and environmental conditions, affect the agricultural yields of the majority of Indian farmers, who tend to own scattered farmland. India loses almost 5.3 billion tons of soil every year on average. Land that is in poor condition can no longer generate enough food. Soil fertility, which is dependent on nutrient levels, is a limiting factor in Indian agriculture; furthermore, soil might be optimal for growing certain crops but disastrous for others. Soil fertility, cultivation strategy, and crop yield predictions may all be informed by the soil's physical, chemical, and biological characteristics. Soil testing, potential agricultural yields, ML, supervised learning, and categorization are all relevant terms.

I. INTRODUCTION

A large portion of India's GDP comes from the agricultural sector. Arable land accounts for 52.8% of India's total farmland, while 4.2% is reserved for permanent crops and 3.5% is used for pastures, according to 2011 statistics. Approximately 42% of the country's workforce is employed in agriculture and allied sectors, which contributed 17.1% to GDP in 2017–18. In the 2013–2014 season, the cultivation areas of key crops were 15 million hectares and 57 million hectares, respectively, according to data from the directorate of economy and statistics (2015). Nowadays, a lot of farmers have their soil tested at the neighboring Krishi Vigyan Kendra (KVK) center to find out what nutrients are present in their soil and in what proportions. Soil testing involves analyzing a soil sample to find out what kinds of nutrients it has, how it is made up, and other interesting facts. It is common practice to conduct tests in order to gauge fertility and identify areas that need improvement. While the Soil Health Card (SHC) does record test

data, it is not immediately useful in determining which crop to plant for optimal yield. Although SHCs are making a difference in soil analysis, most farmers still rely on personal experience and conversations with local experts when deciding which crops to use and how much fertilizer to apply. This method has been used for a long time and is said to offer some advantages. But it's bringing up new, more significant issues, such as soil deterioration from using too many fertilizers, poor yields over time, and negative impacts on both humans and the greater environment. Concurrently, data science and computational science have made tremendous strides. Digitalization is sweeping throughout every industry, including agriculture. There is an ever-increasing amount of satellite images and topography at our fingertips, and the land maps have been digitalized. Soil nutrient composition data in vast quantities is at your fingertips. Additionally, farmers may take use of ubiquitous computing by connecting to the internet and utilizing their mobile phones (for instance, they can use SHC registration applications on their phones).

II. LITERATURE REVIEW

Many people and organizations have worked in this area. There are essentially two primary ways that soils have been classified, and they are based on:

- The soil's biochemical makeup, including factors like temperature, pH, and amount of nitrogen, phosphorus, and potassium, among others.
- Using satellite pictures for remote sensing and assessing the image of the soil

Efforts to study the soil composition fall under the first category, and are described below: "Crop recommendation system for precision agriculture" (S. Pudumalar et al., 2000). [1] makes use of a data mining approach that incorporates information on soil types, soil properties, crop yield, and other site-specific aspects. Because of this, there is less room for error when selecting crops, and

overall output is higher. In order to provide a crop for the site-specific parameters with high accuracy and efficiency, this work proposes an ensemble model with majority voting approach. The model uses Random tree, CHAID, K-Nearest Neighbor, and Naive Bayes as learners. This solves the issue. The article "A Machine Learning Approach to Assess Crop Specific Suitability for Small/Marginal Scale Croplands" was written by Bhimanpallewar R et al. According to the proposed machine learning method in [2], the input parameters are the crop being considered, the environmental factors, and the current availability of soil components. The output is the degree of appropriateness for that particular crop. Because of its limitations, this technique is useful for making judgments like how to make the soil more suitable or how to maintain the area undeveloped for a certain amount of time. According to Senagi K et al. [3], an improved Machine Learning (ML) technique is used in "Using parallel random forest classifier in predicting land suitability for crop production" to determine if land is suitable for growing sorghum crops, given information about the soil's qualities. It arranges trials with the help of PRF, LR, LDA, KNN, GNB, and SVM, or Gaussian Naïve Bayesian. Soil classification systems based on its chemical make-up are still in their infancy. Among these is "An Analysis of Agricultural Soils by using Data Mining Techniques" (Ramesh Babu and Rajesh Reddy, 2018). referenced in [4], [5] by Supriya D., "Analysis of Soil Behaviour and Prediction of Crop Yield using Data Mining Approach," [6] by Sirsat M. et al., and "Crop Recommendation System Using Neural Networks," under the section titled "Using Machine Learning for Land Suitability Classification" [8]. One study that used soil and remote sensing satellite images for research was "Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data" [9] by Kussul N et al., which employed a multi-level DL architecture to classify land cover and crop types from multiple sources and across different time periods in the satellite images. Foundational to the design is a collection of supervised neural networks (NNs), which are used for tasks such as optical picture segmentation and missing data restoration caused by clouds and shadows. Additionally, there is an unsupervised neural network (NN) that serves this purpose. This basic supervised NN architecture compares convolutional NNs (CNNs) with the widely used technique in the remote sensing (RS) community—

random forest—and the classic fully connected multilayer perceptron (MLP). • In their study titled "Improving crop classification with landscape stratification based on MODIS-time series," Driessen B et al. [10] investigate the feasibility of using MODIS images with a moderate resolution for stratification purposes instead of using highly detailed maps of soil and elevation. It makes advantage of land stratification, which involves dividing a monitored region into smaller and smaller pieces, to improve the effectiveness of the monitoring process. On a training set, many classification methods have been used, including RF, SVM, ML, k-NN, and multinomial logistic regression. • A approach that automatically classifies crops from spatio-temporal remote sensing is described in '3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images' [11] by Ji S et al. Article by Rose M. Rustowicz titled "Crop Classification with Multi-Temporal Satellite Images" The use of machine learning to crop categorization using time series satellite images is discussed in [12]. Research into a broader agricultural industry has been attempted and encouraged by the government. A massive quantity of data is being made accessible for study as part of the SHC's broad program. 'Detection & Prediction of Pests/Diseases Using Deep Learning' [13] and similar initiatives follow a similar track.

III. PROPOSED SYSTEM

This opens a new avenue whereby we can carry out the examination of soil, sorting it according to its composition, and then further dividing it into cultivation categories suited to certain crops. Machine learning, big data analysis, and the ever-increasing computing capacity of cloud-based GPU farms will all be used by the suggested system. This s intended to be implemented in two stages by the proposed system: • First stage: system for soil classification based on fertility ratings, nutrients, etc. • In Phase 2, you'll discover the connection between the soil classification groups discovered in Phase 1 and the crop nutrient requirements. This may be achieved by grouping crops that have comparable fertility and soil nutrient needs into labels. Soil analysis is the intended use of the following data by the proposed system: • The soil's biochemical make-up • Pictures of the soil • Data from remote sensing and satellite photography (where available)

Crop nutrient needs (both macro and micro), soil pH, water retention capacity, and electrical conductivity are some of the characteristics that may be used to assign crops to certain soil types. It is intended that the following methods be used in conjunction with the proposed system: • Support vector machines (SVMs), decision trees (DTs), and deep learning using neural networks (NNs) for classification, among other methods. • Bayes distribution, regressions, and other statistical techniques

IV. GAPS AND SUGGESTED SOLUTION

1. So far, every method and previous attempt has relied on either a) analyzing and classifying soil composition or b) using remote sensing satellite photos or soil pictures for analysis and classification. 2. We need to address the issue comprehensively so that we may mix and utilize the best of both methods. Time and money may be greatly saved by automatically classifying soil pictures using previous soil composition categorization. 3. Previous research and solutions have only focused on soil classification or crop recommendation; our goal is to find a more comprehensive and efficient solution that can simplify the whole process and suggest crops that are a good fit by using the soil classification and labeling done in the first phase. 4. The solution uses supervised machine learning (ML) methods to categorize soil and select crops that would be suitable for them. The trained models may be used to detect and classify fresh soil samples, and the same goes for crop suitability. 5. In the end, the proposed method would lead to better agricultural yields and more money in the farmers' pockets.

V. OBJECTIVES

Our goal is to accurately classify the soil using its biochemical makeup and/or the given digital picture. There should be a method to determine the crop's compatibility and achieve the maximum crop production based on the soil labeling. Data science and machine learning are two examples of modern computing techniques that may improve the process's accuracy. up addition to preventing soil erosion and loss, this practice will deter farmers from focusing too much on a single crop and fill up nutritional gaps. The goals that the suggested system aims to achieve are as follows: • Conduct a soil analysis using biochemical and environmental analysis methods, as well as digital images. Soil fertility, nutrient content,

water holding capacity, and other characteristics should be used to categorize the soil. • Suggest which crops might do best in certain types of soil. • Assist in halting soil erosion and preventing soil loss. Reduce fertility by assisting in the adoption of soil-appropriate crops and discouraging over-cultivation of a particular crop. • Increase return on investment while enhancing agricultural productivity.

REFERENCES

- [1] S. Pudumalar, E. Ramanujam, R. Harine Rajashree, C. Kavya, T. Kiruthika, J. Nisha., "Crop recommendation system for precision agriculture," *IEEE Electron Device Lett*, June, 2017. <https://ieeexplore.ieee.org/document/7951740>.
- [2] Bhimanpallewar R et al., "A Machine Learning Approach to Assess Crop Specific Suitability for Small/Marginal Scale Croplands" *International Journal of Applied Engineering Research*, ISSN 0973- 4562 Volume 12, Number 23 (2017) pp. 13966-13973. <http://www.ripublication.com>
- [3] Senagi K et al., "Using parallel random forest classifier in predicting land suitability for crop production", *Journal of Agricultural Informatics* (ISSN 2061-862X) 2017 Vol. 8, No. 3:23-32.
- [4] Ramesh Babu, Rajesh Reddy, "An Analysis of Agricultural Soils by using Data Mining Techniques" *IJESC*, 2017, Volume 7 Issue No.10.
- [5] Supriya D., "Analysis of Soil Behavior and Prediction of Crop Yield using Data Mining Approach", *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 5, Issue 5, May 2017, DOI: 10.15680/IJIRCC.2017. 0505067.
- [6] Sirsat M et al., "Classification of agricultural soil parameters in India" Article in *Computers and Electronics in Agriculture*, April 2017.
- [7] B. Tanmay et al., "Crop Recommendation System Using Neural Networks", *International Research Journal of Engineering and Technology (IRJET)*, Volume: 05 Issue: 05, May-2018.
- [8] M. Mokarram, et al., "Using Machine Learning for Land Suitability Classification".
- [9] Kussul N et al., "Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data" <https://ieeexplore.ieee.org/document/7891032>
- [10] Driessen B et al., "Improving crop classification with landscape stratification based on MODIS-time series", *Laboratory of Geo- Information Science and*

Remote Sensing, Wageningen University and Research Centre, The Netherlands.

[11] Ji S et al., “3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images”, *Remote Sensing*. 2018, 10, 75; doi:10.3390/rs10010075

www.mdpi.com/journal/remotesensing

[12] Rose M. Rustowicz, “Crop Classification with Multi-Temporal Satellite Imagery”, Stanford University, CA, USA.

[13] <https://medium.com/sciforce/machine-learning-in-agricultureapplications-6ab501f4d1b5> and-techniques-

[14]

https://en.wikipedia.org/wiki/Convolutional_neural_network

[15] <https://medium.com/machine-learning-101/chapter-2-svm-supportvector-f0812effc72> machine-theory-

[16] <https://www.tensorflow.org/guide>

[17] <https://www.kdnuggets.com/2017/11/real-world-deep-learning-neuralnetworks-smart-crops.html>

[18] <https://towardsdatascience.com/the-actual-difference-betweenstatistics-and-machine-learning-64b49f07ea3>

[19] <https://www.nature.com/articles/nmeth.4642.pdf>

[20] <http://www.fao.org/3/I9066EN/i9066en.pdf>

[21] <https://soilhealth.dac.gov.in/>

[22]

<https://pib.gov.in/newsite/PrintRelease.aspx?relid=186413>