ISSN:0975-9646
Saurabh M Sirdeshmukh et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 13 (1) , 2022, 39-41

Agriculture Soil Analysis, Classification and Crop Suitability Recommendation Using Machine Learning

Saurabh M Sirdeshmukh, Prof. Prajakta A Satarkar

Computer Science and Engineering Department,
SVERI's College of Engineering, Pandharpur
Punyashlok Ahilyadevi Holkar Solapur University, Solapur
Maharashtra, India.
saurabh.cse2@gmail.com
pasatarkar@coe.sveri.ac.in

***Abstract*— It is very important to increase the crop yield to satisfy the needs of increasing population. Most of the Indian farmers hold fragmented cropland and their yield is dependent on availability of various factors like soil-quality, rainfall, and environmental conditions. Average annual soil loss in India is about 5.3 billion tonnes. Degraded land loses it's capacity to produce adequate yield. Agriculture in India is conditioned by the poor fertility of the soil, which depends on the levels of its nutrients; further a soil can be suitable for certain crops and yield a good production while prove to be otherwise for some other crops. The physical, chemical, and biological properties of the soil are useful to evaluate its fertility, to devise a cultivation plan and to predict the crop productivity.**

***Keywords*— Soil analysis, crop suitability, machine learning, supervised learning, classification.**

I. INTRODUCTION

The agriculture is a backbone of Indian economy. According to data of year 2011, India devotes 60.5% of its land to agriculture, distributed among arable land (52.8%), land for permanent crops (4.2%) and pastures (3.5%). Share of agriculture and related activities was 17.1% of the gross domestic product (GDP) in 2017-18 and it accounts for roughly 42% of total employment in the country. Data from the directorate of economics and statistics (2015) show that in year 2013-2014 the cultivation areas of major crops were 15 and 57 million of hector in Kharif and Rabi seasons, respectively.

Many of the farmers nowadays take their soil samples to nearby Krishi Vigyan Kendra (KVK) centre and get it tested to know available soil nutrients and their proportion. A soil test is the analysis of a soil sample to determine its nutrient content, composition, and other characteristics. Tests are usually performed to measure fertility and indicate deficiencies that need to be remedied. The Soil Health Card (SHC) captures the details of the test but the information they get out of it couldn't help them directly to decide which crop has to be adopted to gain maximum crop yield.

SHCs are certainly helping towards better soil analysis but choosing the crops, fertilizers, and right proportions of their usage is still largely driven by experience and through the discussions with local expert farmers – the overall process is still based on tribal knowledge. This practise is followed since long time and have some benefits. But it is raising newer and more serious problems like soil degradation due to excess quantity of fertilizers, low yield over the time and its harmful effects on human & larger ecosystem.

At the same time, the computational and data science field has rapidly progressed. There is unprecedented amount of digitization happening in all walks of life, including agriculture. The land maps are digitized; there is ever increased satellite imagery and topography available to us. The large amounts of datasets capturing the soil nutrients composition is available. Farmers also have access to ubiquitous computing through mobile phones and network connectivity (for e.g. farmers can register themselves for SHC using mobile apps).

II. LITERATURE REVIEW

There have been various efforts that have gone into this field. Broadly, there have been two main approaches for achieving soil classification, based on:

- Bio-chemical composition of the soil like temperature, pH value, NPK (Nitrogen, Phosphorous, and Potassium) contents etc.
- Leveraging remote sensing satellite imagery and analyzing the soil images

Research efforts mentioned below pertain to first category i.e., examining the soil composition:

- 'Crop recommendation system for precision agriculture' by S. Pudumalar et al. [1] employs the data mining-based technique that uses research data of soil characteristics, soil types, crop yield data

and suggests the farmers the right crop based on their site-specific parameters. This reduces the wrong choice on a crop and increase in productivity. In this paper, this problem is solved by proposing a recommendation system through an ensemble model with majority voting technique using Random tree, CHAID, K-Nearest Neighbour and Naive Bayes as learners to recommend a crop for the site-specific parameters with high accuracy and efficiency.

- 'A Machine Learning Approach to Assess Crop Specific Suitability for Small/Marginal Scale Croplands' by Bhimanpallewar R et al. [2] proposes machine learning technique where input parameters are - existing availability of components in soil, environmental parameters along with selected crop and output is suitability level for respective crop. This system helps to take decisions like how to improve soil suitability or to keep the land free for some duration as it's not capable.
- 'Using parallel random forest classifier in predicting land suitability for crop production' by Senagi K et al. [3] uses an optimized Machine Learning (ML) algorithm for predicting land suitability for crop (sorghum) production, given soil properties information. It sets-up experiments using Parallel Random Forest (PRF), Linear Regression (LR), Linear Discriminant Analysis (LDA), KNN, Gaussian Naïve Bayesian (GNB) and Support Vector Machine (SVM).
- There are few more efforts that classify the soil based on its composition. These include: 'An Analysis of Agricultural Soils by using Data Mining Techniques' by Ramesh Babu, Rajesh Reddy [4], 'Analysis of Soil Behaviour and Prediction of Crop Yield using Data Mining Approach' by Supriya D [5], 'Classification of agricultural soil parameters in India' by Sirsat M et al. [6], 'Crop Recommendation System Using Neural Networks' [7], 'Using Machine Learning for Land Suitability Classification' [8].

Research efforts based on remote sensing satellite imagery and soil images are:

- 'Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data' by Kussul N et al. [9] uses multilevel deep learning (DL) architecture that targets land cover and crop type classification from multi-temporal multi-source satellite imagery. The pillars of the architecture are unsupervised neural network (NN) that is used for optical imagery segmentation and missing data restoration due to clouds and shadows, and an ensemble of supervised NNs. As a basic supervised NN architecture, it uses a traditional fully connected multilayer perceptron (MLP) and the most commonly used approach in remote sensing (RS) community- random forest and compare them with convolutional NNs (CNNs).
- 'Improving crop classification with landscape stratification based on MODIS-time series' by Driessen B et al. [10] tests whether stratification based on moderate resolution MODIS imagery can be used as an alternative to stratification based on detailed soil and elevation maps. It uses the concept of land stratification whereby an area to be monitored is broken up into units to increase the efficiency of monitoring. Classification has been performed using various algorithms (RF, SVM, ML, k-NN and multinomial logistic regression) on a training set.
- '3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images' by Ji S et al. [11] describes a three-dimensional (3D) convolutional neural network (CNN) based method that automatically classifies crops from spatio-temporal remote sensing.
- 'Crop Classification with Multi-Temporal Satellite Imagery' by Rose M. Rustowicz [12] explores the use of time series satellite imagery and machine learning for crop classification.

There have been attempts and encouragement from the government to conduct research into a larger agriculture sector. The SHC is an extensive initiative with large amount of data being made available for analysis. Efforts like 'Detection & Prediction of Pests/Diseases Using Deep Learning' [13] are on similar lines.

III. Proposed System

This opens a new avenue whereby we can carry out the soil analysis, its classification and further categorization into cultivation grades that are suitable for certain crops. The proposed system shall leverage the mature ecosystem and proven infrastructure for machine learning, analyzing large data sets and ever-increasing computation power of cloud-based GPU farms. The proposed system plans to implement this in 2 phases:

- Phase 1: mechanism that classifies the soil according to the fertility grades, nutrients etc.
- Phase 2: find the relation between crop nutrient requirements and soil classification groups found in earlier phase; this can be done by clustering of crops with similar soil nutrient, and fertility requirements into classified soil labels.

The proposed system plans to analyse soil through following data:

- The bio-chemical composition of the soil
- Soil images
- Satellite land imagery and remote sensing data (wherever available and possible)

For mapping the crops to the soil classification groups, we can analyse attributes like:

- The macro and micro nutrients requirements of the crop
- The pH level of the soil
- Water holding capacity & electrical conductivity of the soil

The proposed system plans to use combination of one or more of the following approaches:

- Classification based on decision trees, deep learning using NN (Neural Networks), SVM (Support Vector Machines) etc.
- Statistical tools like – Bayes distribution, regressions

IV. Gaps and Suggested Solution

1. All the approaches and earlier efforts discussed so far were either based on a) soil composition analysis and classification or b) analysis and classification based on remote sensing satellite imagery/soil images.
2. There is need for comprehensive handling of the problem whereby we shall be able to combine and apply best of both approaches. By able to classify the soil images based on earlier soil composition classification, it can reduce a lot of time and cost.
3. Earlier studies and solution have been also limited to either soil classification or crop recommendation; we need an end-to-end streamlined solution that proposes the suitability of the crops that leverages the soil classification/labelling done in the first phase.
4. The solution employs Supervised Machine Learning (ML) techniques for soil classification and crop suitability recommendation, whereby the trained models are helpful to identify and classify new soil samples and crop suitability for the same.
5. Overall, the suggested solution shall thus benefit in increased crop yield and higher financial returns for farmers.

V. Objectives

The objective is to provide the accurate classification of the soil based on biochemical composition and/or digitized imagery provided. Based on the soil labelling, there shall be a way to know the suitability of the crop and thereby have the optimal yield of the crops. Modern computational approaches like machine learning and data science will help lend more accuracy to the process. This will make sure there is no soil loss and degradation, discourages excessive cultivation of a single crop and help improve the missing nutrients in the soil.

The proposed system shall strive to meet objectives below:

- Analyse the soil based on biochemical and environmental composition and/or digitized imagery.
- Classify the soil into appropriate groups based on factors like fertility, nutrients, water holding capacity etc.
- Recommend the suitability of the crops for classified groups of soil.
- Help reduce soil degradation & eliminate soil loss.
- By helping to adopt the crops suitable for the soil, discourage excessive cultivation of a single crop and thereby reduce fertility.
- Improve the crop yield and have higher RoI.

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/IJIRCCE.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-agriculture-applications-and-techniques-6ab501f4d1b5

[14] https://en.wikipedia.org/wiki/Convolutional_neural_network

[15] https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72

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

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

[18] https://towardsdatascience.com/the-actual-difference-between-statistics-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