RESEARCH ARTICLE OPEN ACCESS

# Ensemble Machine Learning Classifier for Soil Nutrient Prediction in Non Agriculture Lands Using Multispectral Remote Sensing Images

Stephen Ajay J[1], Suhaina A[2], Thila Pooja K[3], Dr. V. R. Balaji Ph.D[4].,

Undergraduate Student[1,2,3], Department of Electronics and Communication Engineering, Sri Krishna College of engineering and technology, Coimbatore-641008, Tamil Nadu, India.

Professor[4] , Department of Electronics and Communication Engineering, Sri Krishna College of Engineering and Technology, Coimbatore-641008

## Abstract

Soil Nutrient prediction is essential for precision agriculture on non agriculture lands which termed as residential and forest lands. Soil Nutrient Prediction approaches increase the efficient utilization of available land resources for agriculture and to identify the quality crop to the cultivation in the particular land region. However prediction of soil nutrient in non agriculture lands is carried out using various sensor technologies towards data acquisition. Especially remote sensing technologies acquires the information of the sensing surface in large extent. Acquired information is processed using machine learning techniques which characterizes and explores land surfaces with respect to spectral and spatial information's. In this paper, ensemble machine learning classifier is designed and implemented to process the multispectral images acquired using remote sensing technologies to predict the soil nutrient in the non agriculture lands. Initially, image preprocessing is carried out to eliminate the geometric and radiometric noise in the images. Preprocessed multispectral image is employed to end member extraction technique which is considered as feature for the classifier. Extraction is the process of obtaining the spectral characteristics of the different materials within the mixed pixels and recover the spectrum of the each pure spectral signature. End member of the soil is nitrogen, potassium , phosphorus and organic carbon. Extracted end member is employed to classifier to determine and predict the nutrient level of the soil. Machine Learning classifier employed to the predict the soil nutrient levels are multiple regression , random forest and support vector machine. Classifier model use the ground truth data for nutrient classification in terms of properties of the soil. Experimental analysis of the current ensemble based machine learning classifier is evaluated with cross fold validation in the python framework using Eurosat dataset. Performance of the proposed model proves produces better accuracy compared state of art of approaches.

**Keywords : Soil Nutrient prediction, Remote Sensing technology, Machine learning , EuroSat Dataset, Multispectral Images, Support Vector Machine, Multiple regression, Random Forest**

## 1. Introduction

In recent years, many campaign has been carried out across the world to save the soil. Particularly soil has multiple characteristics as it is suitable several needs of the human beings. Especially Soil has multiple nutrient towards the growth of medicinal crops and food crop across the world. Further soil remove the contaminants from the water. Hence soil is the essential source to human being and their growth. However it becomes vital to determine the nutrient level in the soil in both agriculture and non agriculture land to predict the suitable crop for cultivation and high crop yield[1].

Soil Nutrient monitoring and prediction is essential to farmer for precision agriculture. Identification of soil nutrient in the available lands resources provides solution to determine the suitable crop for cultivation and increases the production of the food grains and medicinal crops. In order to accomplish those task , acquiring the soil parameter become mandatory. It can be obtained using various sensor technologies. In specific, remote sensing technologies can acquires the soil information on large surface. Acquired information is processed using machine learning techniques which characterizes and explores land surfaces with respect to spectral and spatial information's[2].

  

In this paper, ensemble machine learning classifier is designed and implemented to process the multispectral images acquired using remote sensing technologies to predict the soil nutrient in the non agriculture lands. Initially, image preprocessing and end member extraction is established to eliminate the geometric and radiometric noise and to obtain the end member for the classifier. End member of the soil is nitrogen, potassium , phosphorus and organic carbon. Machine Learning classifier employed to the predict the soil nutrient levels are multiple regression , random forest and support vector machine. Classifier model use the ground truth data for nutrient classification in terms of properties of the soil[3].

The rest of the paper is organized as mentioned. In section 2, review of literature related soil nutrient prediction is explored on basis of various functional characteristics. On obtaining those knowledge, section 3 is detailed with proposed design of the machine learning classifier to identify the soil nutrient level on processing the remote sensing images. Experimental results and performance evaluation of the current model against state of art approaches is carried out using eurosat dataset is mentioned in the section 4. Finally article is concluded with its finding and suggestion for research followers.

## 2. Related works

In this section, review of literatures related to soil nutrient analysis using multispectral image is explored and analyzed on its processing steps. Those analysis helps to build the prediction model is summarized and detailed as follows

### 2.2. K nearest Neighbour for Soil Nutrient Classification

In this literature, Machine learning classifier represented as K Nearest Neighbour is employed to predict and classify the soil nutrients on processing its soil features. Soil features of the image is explored using spectral reflectance value. K Nearest Neighbour classifier the soil features into various classes containing soil features of the land surfaces [4]. In addition, it describes the land surfaces in terms of spectral reflectance value of the spectral bands in the spectrum of the particular spatial location of the soil as nitrogen, potassium and phosphorous.

### 2.2. Regression Classifier for Soil Nutrient Classification

In this literature, Machine learning classifier represented as Regression is employed to predict and classify the soil nutrients on processing its soil features. Soil features of the image is explored using spectral reflectance value. Regression the soil features into various classes containing soil features of the land surfaces [5]. In addition, it describes the land surfaces in terms of spectral reflectance value of the spectral bands in the spectrum of the particular spatial location of the soil as nitrogen, potassium and phosphorous.

## 3. Proposed Model

In this section, a ensemble machine learning classifier for soil nutrient prediction is designed for non agriculture land is detailed on basis of preprocessing for noise reduction and end member extraction for classification is as follows

### 3.1. Noise Reduction

Multispectral image is exposed to geometric and radiometric noise on acquisition. Geometric correction and radiometric correction filter is used to eliminate noise in the pixels of the multispectral images[6]. Filter is fabricated to adopt the neighborhood conditions on the selected regions. Filter of the noise is provided as follows

$R[i, j] = \text{Radiometric } F(D_i, D_j)$

### 3.2. End member extraction

Multispectral image contains the abundance of the spectral information representing the micro and macronutrient as contiguous bands in spatial area under analysis. Spectral Reflectance value is computed on each pixel. Unique spectral pure signature is considered as end member of the region. Signatures indistinguishable characteristics of soil on the region of interest[7].

End member of the soil $M=[m_1 \quad m_2 \quad m_3]$

where $m_1, m_2, m_3$ is represented as soil parameter such nitrogen, potassium and phosphorous

  

### 3.3. Soil Nutrient Level Prediction

Soil Nutrient prediction techniques employs the ensemble machine learning classifier to compute the soil nutrient level on basis of soil feature represented as nitrogen, potassium and phosphorous. The classifier employed to categorization of nutrient level is as follows

### 3.1. Random Forest

Random Forest classifier construct the data in the tree structure and it select the random value of the soil features and construct the random vector[8]. Further it calculates the correction of the other feature of the different regions. On basis of strength of the feature , feature map is constructed to represent similar soil nutrient regions using adaboosting technique which combines the similar feature of the particular property. Further it uses the margin function to set the level

Margin function for nutrient discrimination $M_f = P(v)\ ((F(X)_{i,j}+b\ )$

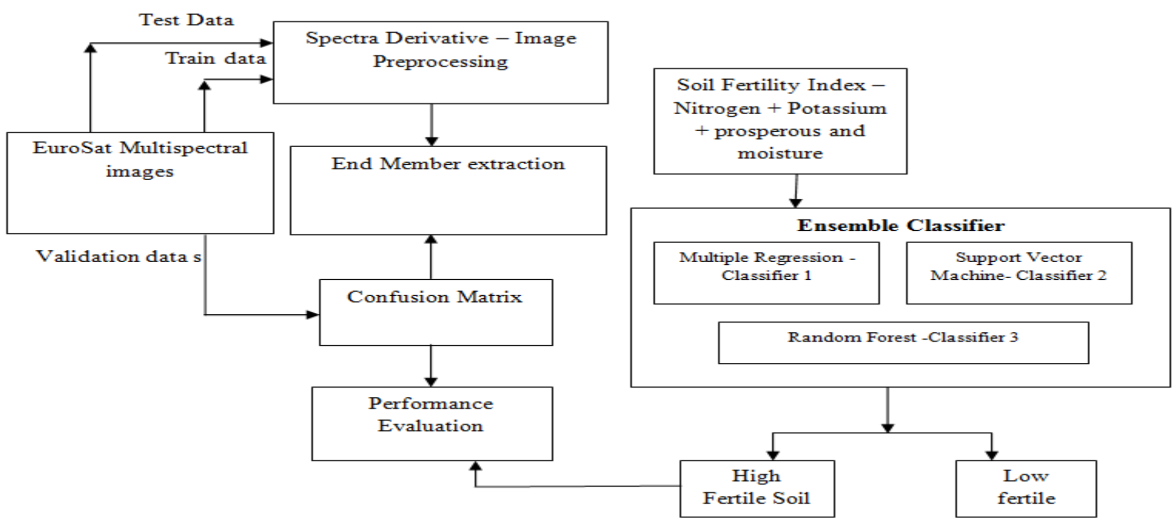


**Figure 1: Architecture Diagram**

### 3.2. Multiple Regression

Multiple Regression works on analyzing the relationship between the dependent feature of the soil nutrients[9]. Dependent feature is represented as criterion and nutrient level is considered as predictor variables.$R^2$ is used to identify the variation of the dependent features with respect to prediction variable. The model for nutrient classification is given as

$$Y' = a + b_1X_1 + b_2X_2 + \ldots b_kX_k$$

x is termed as depended feature and B is regression coefficient

R is considered as Multiple relation coefficient and $R^2$ is considered as coefficient of multiple determination. Figure 1 represents the architecture of the proposed model

### 3.3. Support Vector Machine

Support vector Machine is a classifier employed to classify the soil features of the soil into nutrient level. It uses the hyper plane to process support vector of the soil to classify the soil nutrient. Support vector contain the value of the soil parameters. Maximum Margin separator is computed using the following formula[10]. The soil feature vector with nutrient content is considered as high nutrient and low nutrient.

$$\mathbf{w}.\mathbf{x}^c + b > +1 \quad \textit{for high nutrient}$$

$$\mathbf{w}.\mathbf{x}^c + b < -1 \quad \textit{for low nutrient}$$

  

**Algorithm 1: Ensemble Classifier for Soil Nutrient Classification**

Input: Multispectral images
Output: Nutrient Label L={Fertile L1 or Infertile L2}
Process:
  Preprocess()
    Preprocessed image = noise reduction (Images)
      Spectral reflectance Value = Spectral Signature (Preprocessed image )
  Feature Extraction
    End member EM= Principle Component Analysis ( Spectral reflectance Value )
      End member = Principle Component ={ Nitrogen, Potassium , Phosphorus }
  Nutrient Classification
    Random Forest()
      Tree Structure ( End members)
        Margin = Discrimination ( features )
          Class= feature map
    Multiple Regression()
      Multiple Correlation Coefficient( End Member)
        Class= Nutrient Levels
    Support vector Machine ()
      Margin ( End Member)
        Class= Nutrient Levels

### 3.4. Soil Nutrient

Spectral values of end member of the soil is contains the spectral value for the various band of multispectral images. Table 1 represents the spectral value of the soil parameters.

**Table 1: Spectral Values of the image at 1-100 bands**

| Bands | Organic Matter | pH level | Prosperous | Nitrogen |
|---|---|---|---|---|
| 1-15 | 0.1011 | 0.9028 | 2.7771 | 0.0591 |
| 15-30 | -0.5156 | 0.9021 | 2.7855 | 0.0625 |
| 30-45 | 0.1896 | 0.9020 | 2.7858 | 0.0626 |
| 45-60 | 0.2514 | 0.9010 | 2.8052 | 0.0704 |
| 60-75 | 0.5205 | 0.9007 | 2.8104 | 0.0725 |
| 75-90 | 0.4254 | 0.9005 | 2.8138 | 0.0739 |
| 90-105 | -0.6283 | 0.9014 | 2.7975 | 0.0673 |

## 4. Experimental Results

Experimental of the current ensemble approach is carried out on the EuroSat dataset[10] and performance of the current approach is evaluated using the various performance measures such as precision, recall and F-measure to evaluate the efficiency and accuracy against the conventional approaches.

  

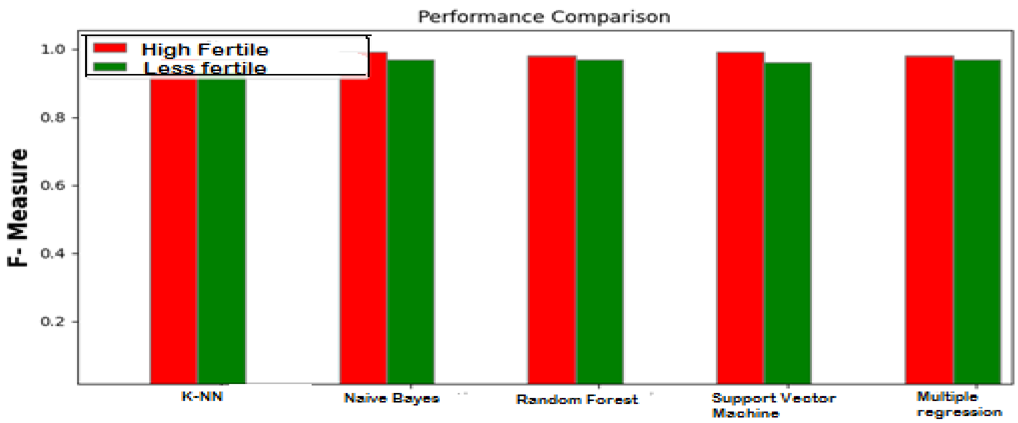


**Figure 2: Performance evaluations of the ensemble machine learning classifier in soil nutrient prediction**

Proposed approach generate high Performance as it incorporates end member extraction analysis for spectral signature to classify soil nutrients of the soil parameter compared to existing approaches. Prediction accuracy is computed and it is depicted in the Figure 2 and Table 2.

**Table 2: Performance Evaluation**

| **Technique** | **High nutrient** | **Low Nutrient** |
|---|---|---|
| K Nearest Neighbour - Existing 1 | 0.89 | 0.90 |
| Regression Analysis –Existing 1 | 0.91 | 0.94 |
| Random Forest –Proposed | 0.87 | 0.97 |
| Multiple regression -Proposed Classifier | 0.85 | 0.94 |
| Support Vector Machine- Proposed Classifier | 0.88 | 0.92 |

As the proposed approach possessed higher accuracy when compared with the conventional approach, we opted for the proposed approach composed of multiple classifier as ensemble model to assign nutrient label to soil of the non agriculture land.

## Conclusion

We designed and implemented a ensemble machine learning architecture encomposing of Support vector machine, Random Forest and Multiple Regression to predict the soil nutrient content in in the non agriculture land using multiple spectral images. Particular model initialized with noise reduction and end member extraction process. The spectral signature of the model is extracted after noise reduction. Particular end member of the soil is extracted using principle component analysis. Extracted end member is classified into high nutrient and low nutrient classes of the soil. On experimental analysis, it is considered to more accurate compared to the conventional approaches. It is required to incorporate the constraints to reduce the over fitting issues is considered as future direction of the particular work.

## References

1. Kamal Das, J Singh, J Hazra, and S Kalyanaraman, "Evaluation of land surface model against smap and in-situ observations for indian region, " in IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2018.

  

2. Aniket Chakravorty, Bhagu Ram Chahar, Om Prakash Sharma, and CT Dhanya, "A regional scale performance evaluation of smos and esa-cci soil moisture products over india with simulated soil moisture from merraland, " Remote Sensing of Environment, 2016.

3. Gunkirat Kaur , Kamal Das, Jagabondhu Hazra , "Soil nutrients prediction using remote sensing data In Western India: An Evaluation Of Machine Learning Models" ,2020.

4. Gerald Forkuor, Ozias K. L. Hounkpatin, Gerhard Welp, Michael Thiel, " High Resolution Mapping of Soil Properties Using Remote Sensing Variables in South-Western Burkina Faso: A Comparison of Machine Learning and Multiple Linear Regression Models" ,2017

5. Kamal Das, Jitendra Singh, Jagabondhu Hazra, Shivkumar Kalyanaraman, " Evaluation of land surface model against SMAP and In-situ observations for Indian region"-2017

6. Kaihua Liao , Shaohui Xu , Jichun Wu & Qing Zhu, " Soil Science and Plant Nutrition"-2013

7. Yufeng Ge , J. A. Thomasson , Ruixiu Sui, " Remote Sensing of Soil Properties in Precision Agriculture: A Review"-2016 28

8. Ccile Gomez, Raphael Viscarra Rossel, and Alex Mcbratney, "Soil organic carbon prediction by hyper-spectral remote sensing and field vis-nir spectroscopy:An australian case study, " - 2008

9. Kaihua Liao, Shaohui Xu, Jichun Wu, and Qing Zhu, "Spatial estimation of surface soil texture using remote sensing data, " Soil Science and Plant Nutrition, vol. 59,08 2019.

10. Gerald Forkuor, Ozias Hounkpatin, Gerhard Welp, and Michael Thiel,"High resolution mapping of soil properties using remote sensing variables in south-western burkina faso: A comparison of machine learning and multiple linear regression models " PLOS ONE, 2017.

  