IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727,
PP 83-91
www.iosrjournals.org

Remote Sensing and Data Mining Techniques Applied on Soil Characteristics Data Classification

Miss. Bhise Pratibha.R[1], Dr.Kulkarni Sonali.B[2]
[1](Department of Computer Science and Information Technology, Dr.BAMUniversity, Aurangabad, India)
[2](Department of Computer Science and Information Technology, Dr.BAMUniversity, Aurangabad, India)

Abstract: *Recently we see the rapid development in Remote Sensing and Data Mining .It has free various real life applications in several domains such as Urbanization &transportation, National Resource Management, National Security and Agriculture is one of them. This review highlights the precedence of Remote Sensing, and Data Mining techniques for examining the soil properties .As we know soil is a vital part of the environment, survival and humanistic welfare .Soil is formed by minerals, organic matter and living organism components. Soil prediction is dependent on the soil substances. Soil physical, Biological and chemical properties plays a very significant role in agricultural field. Accelerated low cost and predictable assessment of soil quality under agricultural management is necessary to accomplish convenient observation of the effects of various management practices on soil conditions to avoid soil degradation and ensure feasible soil productivity and also soil security. Here we represent all remote sensing and data mining techniques for analysis of soil characteristics and data classification.*

Keywords: *Data Mining, Heavy Metals, Remote Sensing, VIS/NIR*

I. Introduction

Remote Sensing is nothing but the acquire information about earth surface features without physical contact with it. Whereas Data Mining is a study of investigating extensive previous available databases in order to propagate new information. Observing soil quality is a vital question for a sustainable evolution of economy. Soil is a natural resource it is linked with human life .The definition of Soil is a living, breathing, naturally occurring dynamic system at the interface of air and rock. Soil forms in response to forces of climate and organisms that act on parent material in a specific landscape (topography) over a period of time. soil quality presenters are Physical, Biological and Chemical properties. Which are used to monitor advances in soil. Soil quality presenters are important to evaluate soil management practices and techniques, it predicts the future valu8e of soils for agricultural, forestry and also describe soil quality to that of other resources. Soil is composed of minerals (Sand, Silt and Clay), water, air, organic matter, and countless organisms that are the decaying remains of once-living things. Soil is a critical part of successful agriculture and is the original source of the nutrients that we use to grow crops. Soil is having need of Macronutrients& Micronutrients for crop production. The nutrients required by the plants in substantial volume. These are Carbon, Hydrogen, Oxygen, Nitrogen, Phosphorus, Sulphur, Calcium and Potassium are obtained from air sodium are present which are required by higher plants .A lack of any one of the micronutrients in the soil can limit growth, even when all other nutrients are present in adequate amounts.[1]So to predict available nutrients percentage or also efficiency of nutrients we required some classification techniques which are available in remote sensing ,GIS and data mining. Remote Sensing and GIS gives us hyper spectral non-Imaging spectroscopy technique which is low cost ,time consuming ,fast ,and easy to use .It can be used to resolve soil nutrient status ,available soil balance property, maintain soil fertility ,decrease nutrient loss, and also helps in to reduce the environmental pollution. The ultraviolet (UV), visible (VIS), near infrared reflectance (NIR), and mid infrared (MIR) spectroscopy techniques used to determine the phosphorus in soil samples .New various approaches in spectroscopy have provided modern approaches to resolve the absorption of components in a discrete scope of agricultural products.[2]Various algorithms are present for the measures of near infrared spectra in soil study .It is successfully predict the Carbon and Nitrogen derived from various sources in soil .Reflectance wavelength 400-1300 nm is sensitive for organic C. In this they used VIS-NIR Spectroscopy as an alternative technique to predicting the soil aggregate stability indexes. It is similar to the some basic properties of soil. It is study of various soil types which are correlated to organic carbon content for soil aggregate stability indexes. The utility and applicability of Lab VIS-NIR Spectroscopy as an alternative method for soil aggregate stability indexes estimates.[3]They suggested that the soil of the oasis-desert ecosystem has low soil water, high concentrations of salt and a structuralizes soil with very low organic matter, but concentrations of soil nutrients increased and improvement occurred after land use shifts from natural land uses to cultivation.. The four natural lands have distinct soil characteristics, which should be considered during the process of soil reclamation.[4]In this study

they introduces a tractor-driven me reflectance spectroscopy (Vis–NIRS) encounter preliminary and carefulness agriculture requirement , where the SOC variability among the experimental plots was pointedly induced.[5]

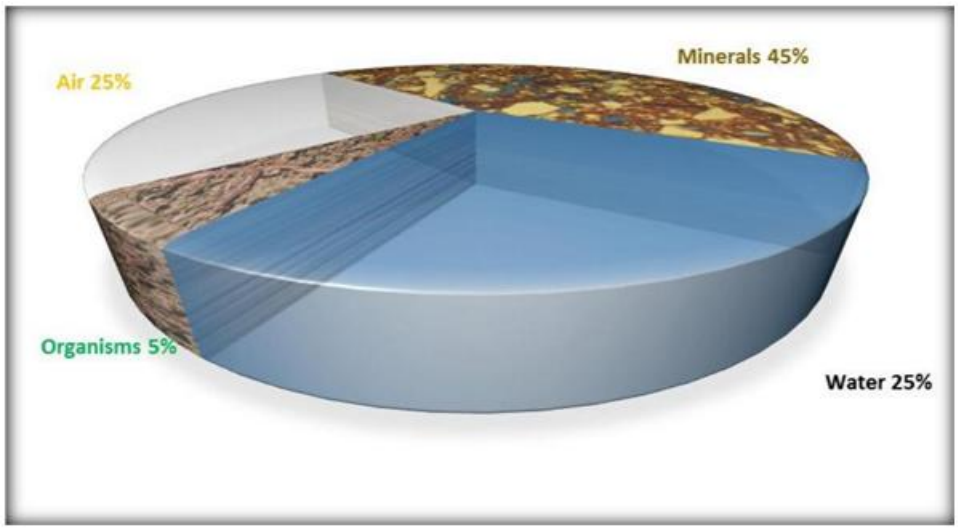

Fig1: Soil Profile

This study underscores the potential application of reflectance spectroscopy as a reliable diagnostic screening tool for assessing soil quality. Classification of soils into spectrally characterizes substances produce a support for partially accurate and denotative interpretation for developing SSQI. This study proposes a framework for the use of VIS–NIR–SWIR spectroscopy as a tool for assessing soil quality. [6]The carbon (C) and nitrogen (N) in soils are crucial to a number of natural processes related to soil health and fertility. Understanding of the processes which are responsible for soil carbon and nitrogen changes with respect to land-use change and land management activities requires frequent measurement The Evaluation potential of near-infrared reflectance (NIR) spectroscopy used to resolve the carbon and nitrogen content in soils and also the efficacy of NIR spectroscopy on the prediction of carbon and nitrogen content in Indian soils is highly reliable or Water content in soil samples could affect the NIR absorbance spectra and it affects the quantification of carbon and nitrogen.[7]This study evaluated the efficiency of VIS–NIR spectroscopy and chemo metric analysis for rapid and cost-property evaluation of soil quality indices. VIS–NIR spectroscopy can be used as a reproducible approach for observing the goods of agricultural management on soil quality and there is no good reason to progress to the development of field instrument methods to permit large area, high density monitoring to address soil threats related to agricultural production. and also reported results of Carbon(C),Moisture content(MC) ,pH and Phosphorous.[8]They examine the ability of reflectance spectroscopy to predict some of the most important soil specification for irrigation such as field capacity (FC), wilting point (WP), clay, sand, and silt content. The raw reflectance data gave better results for FC and WP; while first order derivative data gave They select the wavelengths according to the highest positive and negative correlations between soil property and raw reflectance or first order derivatives. predicting soil properties using spectral reflectance data .Aim of this was the models which are developed by infrared spectroscopy could be use to evaluate organic carbon content(C) ,total nitrogen content (N) and the C:N ratio in the particulate organic matter (POM) and particle size fraction samples of Brookston clay loam. So that the results demonstrate that both MIR and NIR techniques can be used as different techniques for supposing organic C and total N in the POM and particle size fractions of soil elements . Anyhow ,the NIR model is more appropriate for supposing organic C and N in POM ad sand fractions than the MIR model ,When the MIR model is exclusive to the NIR model for estimating organic C in silt and clay fractions and N in clay fractions.[10]This study estimated an in-field near infrared (NIR)instrument to forecast the contents of total nitrogen (TN) ,organic carbon (OC) ,potassium (K), sulphur (S), phosphorus (P), pH, and electric conductivity (EC) in soil vineyard data . Soil data were study estimated an in-field near infrared (NIR) instrument to forecast the contents of total nitrogen (TN) ,organic carbon (OC) ,potassium (K), sulphur (S), phosphorus (P), pH, and electric conductivity (EC) in soil vineyard data . Soil data were investigated using a portable NIR spectrophotometer (ASD Field Spec III, 350–1800 nm). [11]As we see Mid-infrared diffuse reflectance spectroscopy can provide very rapid, cheap and approximately specific prognosis for a number of soil characteristics . Maximum studies have established that it is probable to evaluate chemical properties which are related to surface and solid material composition. They concluded that the physical and mechanical properties can be related back to the fundamental soil properties such as clay content,

carbon content, cat ion exchange capacity and bulk density and also NIR spectroscopy used to predict Nitrogen(N),Phosphorus(P),Potassium(K),Organic Matter(OM) and pH content in a loamy mixed soil data.[12]

1.1. Accessing Reflectance Information From Soil

To acquire reflectance information from the VNIR-SWIR region, several sensors, methods, protocols and platforms are used. Basically, the VNIR-SWIR sensors contained tools these are filters, grating, prisms, interferometer, etc. they divide the measured radiation into specific wavelengths, certain detectors which are sensitive to various regions, and various other optional optical components such as lenses, collimators, beam splitters, integrating spheres, optical fibers, etc... The radiation source is particularly us as a tungsten halogen light bulb at the time of recording spectra under laboratory conditions, and the sun when recording spectra in the field. These sensors are mostly regularly calibrated to provide the radiance values but in practice, the reflectance is calculated against a white reference target (Spectral on) that is assumed to reflect 100% of the radiation in the sample geometric configuration. There are two types of sensors one is point and another is image sensors. The point sensor records a single spectrum of the target with a line-array detector although averaging the mixed information of all components within the sensor's field of view. The size of the single pixel being measured depends upon the optics and the distance from the target. The image sensor provides data similar to the point sensor except that an area-array detector is used that enables the acquisition of spectral-cube information constructed from spatial and spectral views of the area (target) in question. Image sensors usually have lower spectral resolution as well as lower signal-to-noise ratio, as the integration time over the targets is less than in the point spectrometer measurements. Laboratory sensors are more accurate, with higher spectral resolution and a better signal-to-noise ratio: the integration time is long and the geometric scheme is constant. They usually include a sample holder and a stable light source. Field sensors are generally battery-operated and are less accurate than laboratory sensors. [8][13]

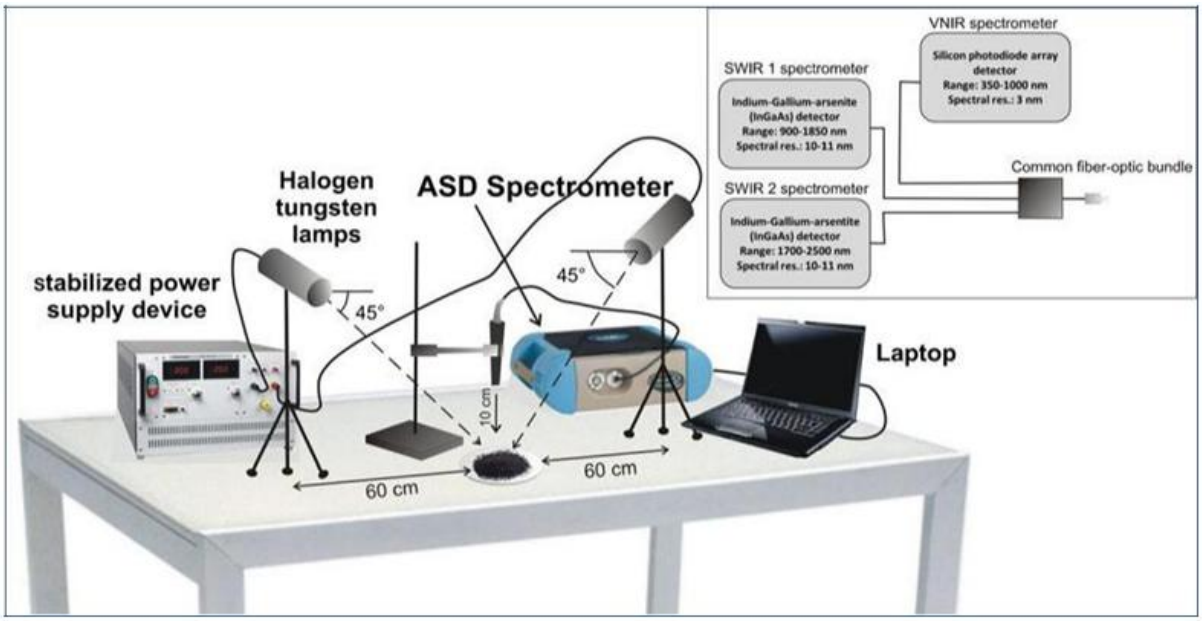

Fig2: Working of ASD Spectrometer

II. Reflectance Spectroscopy To Achieve Heavy Metals In Soil

To monitor the enlargement of heavy metals effectively and escape the illness to the health of agricultural soils, a favorable access is to presume low concentrations of heavy metals in soils using visible and near infrared (VNIR) reflectance spectroscopy coupled with calibration techniques. The objective of this is compare the performance of a combination of partial least squares regression with genetic algorithm (GA-PLSR) against general PLSR for predicting low concentrations of four heavy metals (i.e., As, Pb, Zn and Cu) in agricultural soils.[14]The distinguishing attribute of the original reflectance data were correlated with models borrowed from first and second derivatives of the reflectance data. The results indicate that the models copied from the first derivative of the reflectance data evaluate heavy metals extremely more precisely than model derived from the original reflectance. It was also found in this study that there is no need to use second derivative for modeling heavy metal soil contamination. Finally, the results indicate that assessment were of greater efficiency for arsenic and lead compared to other heavy metals, while the assessment for silver was found to be the most specious.[15] Reflectance spectroscopy Visible –Near Infrared (VIS-NIR)is used to

estimate soil heavy metals such as concentration of Cobalt present in certain samples of soil .[16]also VNIR ranges are used to evaluate Content of arsenic in soil fragment.[17] For recognition of content value of arsenic in soil samples we used Normalized Difference Spectral Index .[18][22]NIRS and DRIFT are useful to successfully evaluate Iron, Cd, Cu, Ni, and Zn .[19]This discovery is consistent with Cr ,Cu Zn and As have important correlations with Fe2O3,Al2O3 and TOC but Cd ,Pb and Hg only display a powerful correlation with TOC**.**[20] Concentration for As,Cd,Cu,Fe,Hg ,Pb,S,Sb and Zn predicted using VNIR.[21]Reflectance Spectroscopy is a tool to study Concentration of Iron (Fe) ,As and Cu in soil fragments from the field .[23] Range of reflectance spectroscopy (400-2500) estimate the impurity for Mn ,Pb, and Zn for Ni and Cr but not give the result for Fe ,Cu ,Cd,Ec,and pH.[24]

III. Quantitative Measurement Of Spectroscopic Data

The data which is collected by spectroscopy called as Spectroscopic data. It is present in multivariate in nature. There are two approaches are available to study spectroscopic data those are

1) Supervised, and
2) Unsupervised

Therefore the first approach is valid for the IR region by using selected wavebands and pre-treatment to inhibit saturation, the second represents the case of reflectance spectroscopy across the VNIR-SWIR region. For that purpose, multivariate statistical techniques (also called chemo metrics) are required to abstract the information about the quality attributes that is invisible within the spectral information. Especially, this involves regression techniques coupled with spectral pre-processing.

3.1 Preprocessing

Spectral pre-processing techniques are used to extract any inappropriate information which cannot be handled properly by the modeling techniques. The pre-processing techniques include averaging, centering, smoothing, standardization, normalization and transformations, among others. Some of the most common pre-processing techniques are presented here.

3.2 PLSR (Partial Least Square Regression)

Introduced in 1983 by Wold et al., partial least squares regression (PLS) is similar to PCR, but in PLS the PCs are constructed such that they include the chemical reference (Y variables, dependent data) in the calculation process. This technique orders the PCs according to their relevance for predicting the dependent variables, rather than to their description of the most variance of the spectral data. This method excels when the For the PLS analysis the wavelengths contributing most to the prediction model are found in the VIS (maximum at 650 nm), around the water absorption features (1400 nm and 1900 nm), and in the SWIR region (2212 nm). The optimal balance between model complexity and predictive power was found for 2 latent variables. Dependent data (X variables) express common information, as usually happens in spectral data. The required number of PCs is typically smaller than that in a PCR calibration model for similar model performance [16]][22] [24][25][26][27]

3.3 MSC (Multiplicative Scatter Correction)

Multiplicative scatter correction (MSC) is one of the most commonly used normalization techniques. In MSC, the light scattering is estimated for each sample relative to an ideal sample obtained by averaging the complete wavelength range of the data set. Each spectrum is then corrected such that all samples appear to have the same scatter level as the reference spectrum.[13]

3.4 Continuum Removal

The continuum is the background absorption on which other absorption features are superimposed. The spectrum is divided by a convex hull that is fitted over it. Straight-line segments were fitted to these points and the continuum-removed (CR) spectrum was calculated by dividing the original reflectance values by the corresponding values of the continuum line. [26] [27]

3.5 Derivatives

Derivation is often used to remove baseline shifts and superposed peaks. Second-derivative spectra can correct for both additive and multiplicative effects (like MSC). Takes the first derivative of the data. The algorithm uses a specified gap distance to skip that number of points to take the differences instead of adjacent data point. They are usually calculated according to the Savitzky-Golay algorithm the parameters of the algorithm (interval width, polynomial order) should be carefully selected to avoid amplification of spectral noise. [10][29]

3.6 Spectral Mixture Analysis

Spectral mixture analysis (SMA) is a widely used method to determine the sub-pixel materials that fundamentally contribute to the spectral signal of mixed pixels. This is of particular importance for obtaining quantitative estimates of distinct materials, a typical application of remote sensing hyper spectral data. SMA aims to decompose the measured reflectance spectrum of each pixel into the proportional spectral contribution of so-called end members (EMs). In recent years, many authors have proposed and used a more complex model, in which both the number and the set of EMs vary dynamically on a per-pixel basis; this has become known as multiple EM SMA (MESMA). The idea consists of restricting the large set of possible EMs to a small set of better suited EMs, which can be different for each pixel, thereby allowing an a EMs. An improved strategy, called variable MESMA (VMESMA), allows segmentation of the image to increase flexibility and accuracy curate decomposition using a virtually unlimited number of EMs. An improved strategy, called variable MESMA (VMESMA), allows segmentation of the image to increase flexibility and accuracy.[13]

3.7 Modeling

Modeling refers to relating a set of spectral parameters that are derived from the spectral information (before or after the aforementioned pre-processing treatment), to the real chemistry of the material in question. This is done by using a set of well-known samples as a training group. The data are divided into three groups: training, validation and test. The relationship between the chemistry and the spectroscopy data is found via the training group and simultaneously cross-validated by the validation group. Finally, the model is applied to the test group, independent of the training and validation process. Multivariate regression techniques are modeling methods that search for the relationship between two matrices: the spectral data matrix that can be very complex due to large amounts of data (X variables, the independent data), and a specific chemical reference value data matrix (Y variables, the dependent data).[10][28]

3.8 MLR (Multiple Linear Regressions)

Multiple linear regressions (MLR) are a classical method that creates a linear combination of the spectral values at every single wavelength to correlate as closely as possible to the dependent reference values. The regression coefficients are estimated by minimizing the error between predicted and observed response values in a least squares sense. MLR models typically do not perform well with spectral data because spectral data usually exhibit high co-linearity, noise and more variables, i.e. more spectral bands, than measured samples.[21]

3.9 PCR (Principal Component Regressions)

Principal component regression (PCR) is a combination of principle component analysis (PCA) and MLR. The independent data matrix (spectral data) is transformed by PCA, and the first few principal components (PCs), which represent most of the independent data variance, are used as inputs for the MLR model instead of the original spectral data. The advantage over standard MLR is that PCs are uncorrelated, and the noise is filtered. The first few PCs are usually sufficient for a robust model and over-fitting issues can be eliminated. Although PCR consists of the two most studied multivariate methods (PCA and MLR), the major criticism against it is that the PCs (several first PCs) selected for the MLR input are not necessarily the best predictors for the reference data. There is no guarantee that the first PCs will include the spectral data related to the specific dependent variable that needs to be modeled.[10]

3.10 ANN (Artificial Neural Networks)

These are based on their ability to "learn" during a training process where they are presented with inputs and a set of expected outputs. The basic structure of an ANN consists of three layers of "nodes" or "neurons": an input layer (i.e. spectral data or PCs), a hidden layer (which can consist of one of more nodes), and an output layer (which combines the outputs of the hidden layer into a single output). The node is a fundamental processing unit; each node has a series of weighted inputs, either from an external source or the output from other nodes. The inputs to the node are analogous to synapses, and the weights correspond to the strength of the synaptic connection. The sum of the weighted inputs is transformed with linear or nonlinear transfer functions, a popular nonlinear transformation function being the sigmoid function. The learning (or training) is an iterative process in which the resultant output is compared to the provided expected output, and an algorithm adjusts the weights accordingly. This method was first tested in the field of spectroscopy on simulated data. [21]

3.11 Stepwise Regression Results

Stepwise regression analysis methods commonly used to identify the wavebands sensitive to a certain chemical constituent, and to demonstrate these wavebands has a good correlation with the concentration of a

certain chemical constituent. Accordingly, we can use these determined locations of the wavelength (band values) to estimate the concentration of a certain chemical composition. However, there are two aspects of deficiency: firstly, there exists over fitting phenomenon in establishment of regression model. This phenomenon mainly appears while the sample size is less than the amount of wavebands. Then spectral reflectance values may not correlated with certain chemical composition while its noise pattern may be related to certain chemical composition.[29]

IV. Analysis Of Soil Elements Using Data Mining

Data mining is the technique to disclose impressive knowledge from huge amounts of data. It is the procedure which gives its results in exploration of new patterns in vast data set. The important objective of the data mining procedure is to abstract knowledge from a current data set and reconstructs it into a people intelligible evolution for further use. Data mining work can be classified into two categories:

1] Descriptive data mining and 2] Predictive data mining.

Descriptive data mining tasks represent the general characteristics of the data in the database although Predictive data mining is used to describe precise value based on patterns resolved from well-known results. As we know soil is a crucial element. Beyond it we wouldn´t be capable to grow plants, which are used as basic need or food for both humans and animals for living life. Most of the peoples are depends on agriculture for its livelihood. It Means agricultural development are increase day by day .Quantity of data are increased means it requires spontaneous techniques to extract the useful information when it needed. As we see very less farmers are known or using the new approaches, devices, tools, techniques of framing for prominent production. A very few farmers are actually using the new methods, tools and technique of farming for better production. Data mining can be used for predicting the future trends of agricultural processes. So, for the predicting future trends we have need for some classification techniques which are presented in below.

4.1 Bayesian classification is used for soil texture classification, it measures the relationship between soil science and data mining classification technique .These classifiers are statistical classifiers. It is used to conclude class membership probabilities like just as the probability which given tuple belongs to a particular class .the classification which used in this that is naïve Bayesian Classifier .It is equivalent in performance with decision tree and other selected classifiers such as neural network. [30]

4.2 This Rule-Based classification technique is used to classify soil texture such as coarse-grained soils (2) fine-grained soils and (3) highly organic soils. This technique represented as a set of IF-THEN rules .In this we firstly examine how these rules are used for classification then also study the approaches in which they can be generated that is either from a decision tree or precisely from the training data using a sequential covering algorithm.[35]

4.3 The decision tree algorithms consist of Genetic Algorithm, The decision tree induction is the study of decision trees from class –labeled training tuples, It is a just like as flowchart tree structure, where each internal node means non –leaf node which represent a test on an attribute, each branch denotes an result of the test, and each leaf node or terminal node control a class label. **Genetic Algorithm** is a method used for expansion and also it called research technique which uses techniques based on biological transformation .such as mutation, selection and inheritance .There are variety of applications in which genetic algorithms have been widely used for development and knowledge diagnosis. such as agriculture .In agriculture it is used for are used to investigate soil characteristics such as Depth, Sand, Silt, Clay, Sandbysilt, Sandbyclay, Sandbysiltclay, Texture Class .GA gives some unique features other than other tree algorithms and also it gives prominent results on various difficult problems.[40]

4.4 The technique Fuzzy Classification Rules are used to represent soil chemical properties like pH.,EC,it is a technique which was proposed to represent management of data and information preparation ,non statistical unpredictability. It is statistically used to represent unpredictability, ambiguity and also provide prescribe tools for dealing with the impression intrinsic to various problems. [36][40]

4.5 The unsupervised classification algorithm Fuzzy C – Means are used to evaluate soil macro nutrients and soil micronutrients like Phosphorous ,Potassium, Iron, Zinc, Manganese, Copper, Boron. In this technique class labels are unknown. It is used to deal unsupervised data with unpredictability. the main motive of this algorithm is to group the objects into clusters which are based on their observable features in that each cluster contains objects which contribute some of the important properties.[31][32][33]In data mining domain various techniques or decision algorithms are available to use for solving different problems related to soil data

classification. These techniques and decision algorithm are discussed by various researchers are presented at below in Table 1.

Table01. Different Techniques Used By Different Authors For Soil Data Analysis

Sr.No	Author, Year & Reference ID	Data Mining Classification Technique, Algorithm
1	PremKumar Chandrakar, Sanjay Kumar and Dewashish Mukherjee,2011 [30]	Bayesian Logistic Regression, ayes HNB, Naïve Bayes,Tree:classifiers.trees.Id3,classifiers.trees.NBTree,classifiers.trees.J48,classifiers.trees.Decision,classifiers.trees.BFTree,classifiers.trees.RandomTree,andRule:classifiers.rules
2	P.Bhargavi,Dr. S. Jyothi [31]2011b	GATree, Fuzzy Classification Rules, Fuzzy C – Means Algorithm
3	Jay Gholap, Anurag Ingole, Jayesh Gohil, Shailesh Gargade, Vahida Attar,2012a [32]	Naive Bayes, J48 (C4.5), JRip
4	Jay Gholap,2012b [33]	J48 (C4.5),NBTree,SimpleCart
5	Rashmi Arora, Savneet Kaur, Shweta Taneja,2012c [34]	K-Means algorithm
6	Vrushali Bhuyar,2014a [35]	J48 decision tree classifier, Naïve Bayesian classifier, Random forest
7	Velide Phani kumar , Lakshmi Velide,2014b [36]	Naïve Bayes,J48 (C4.5), JRip
8	M.C.S.Geetha,2015a [37]	Association Rule Mining-Apriori Algorithm, Clustering, Regression, Classification
9	Dr. S.Hari Ganesh, MrsJayasudha,2015b [38]	Naive Bayes, JRip, J48(C4.5)
10	M.Suriyapraba,S.Sasikala, 2016 [39]	NB Tree, Simple Cart, Navie Bayes, GA Tree, Fuzzy Classification, Fuzzy C-Means Algorithm, J48, JRip

V. Conclusion

Day by day the need to maintain the health of the soil resource base as an important for sustainable development is progressively being identified. Science and technological developments in Remote Sensing and Data Mining are producing the new opportunities for low cost and efficient applications for characterizing and monitoring the health of the soil resource base. It is important that we learn more about soil to sustainably maintain and protect it for future generations. The Remote Sensing gives us various hyper spectral privileges to do rapid and cost effective analysis of soils that could provide new opportunities and also various different quantitative measurement techniques for farmers, land managers, local authorities and researchers in assessing and managing soil quality. Soil is a critical part of successful agriculture and is the original source of the nutrients that we use to grow crops. Nutrients are also a part of the food of animals. At the end, we benefit from healthy soil. The healthiest soils produce the healthiest and most abundant food supplies. According to the literature here we concluded that Data mining and Remote Sensing domain provides various algorithms and prognosis techniques for classification of soil characteristics data.

References

[1]. http://www.cropnutrition.com/efficient-fertilizer-use-guide

[2]. A. Morón & D. Cozzolino," Measurement of Phosphorus in Soils by Near Infrared Reflectance Spectroscopy: Effect of Reference Method on Calibration", *Communications in Soil Science and Plant Analysis,2007, ISSN 0010-3624 print/1532-2416.*

[3]. C. Gomez , Y. Le Bissonnais , M. Annabi , H. Bahri , D. Raclot,"Laboratory Vis–NIR spectroscopy as an alternative method for estimating the soil aggregate stability indexes of Mediterranean soils",*ELSEVIER,Geoderma 209–210 (2013) 86–97.*

[4]. Lu Gong , Qiyang Ran , Guixiang He , Tashpolat Tiyip ,"A soil quality assessment under different land use types in Keriya river basin, Southern Xinjiang, China" ,*Soil & Tillage Research 146 (2015) 223–229.*

[5]. Andrei Rodionov , Gerhard Welp , Lutz Damerowb, Toni Berg , Wulf Amelung , Stefan Pätzold ," Towards on-the-go field assessment of soil organic carbon using Vis–NIR diffuse reflectance spectroscopy: Developing and testing a novel tractor-driven measuring chamber",*ELSEVIER,2014,Soil & Tillage Research 145 (2015) 93–102.*

[6]. Tarin Paz-Kagan, Moshe Shachak, Eli Zaady, Arnon Karnieli,"A spectral soil quality index (SSQI) for characterizing soil function in areas of changed land use", *ELSEVIER, Geoderma, 2014, 230–231 (2014) 171– 184*

[7]. J. Dinakaran, Ankita Bidalia, Amit Kumar, Mohammad Hanief, Archana Meena & K.S. Rao," Near Infrared-Spectroscopy for Determination of Carbon and Nitrogen in Indian Soils", *Communications in Soil Science and Plant Analysis,2016,ISSN: 0010-3624 (Print) 1532-2416.*

[8]. A.M. Mouazen , M.R. Maleki, J. De Baerdemaeker, H. Ramon," On-line measurement of some selected soil properties using a VIS–NIR sensor",*ELSEVIER,2007, Soil & Tillage Research 93 (2007) 13–27*

[9]. Hakan Arslan & Mehmet Tasan & Demet Yildirim &Eyüp Selim Koksal & Bilal Cemek," Predicting field capacity, wilting point, and the other physical properties of soils using hyper spectral reflectance spectroscopy: two different statistical approaches",*Springer,2014, Environ Monit Assess (2014) 186:5077–5088*.

[10]. X . M. Yang, H. T. Xie , C. F. Drury , W. D. Reynolds, J. Y. Yang & X. D. Zhang,"Determination of organic carbon and nitrogen in particulate organic matter and particle size fractions of Brookston clay loam soil using infrared spectroscopy", *European Journal of Soil Scienc,2012, European Journal of Soil Science, April 2012, 63, 177–188*

[11]. D. Cozzolino , W. U. Cynkar , R. G. Dambergs , N. Shah & P. Smith," In Situ Measurement of Soil Chemical Composition by Near-Infrared Spectroscopy: A Tool Toward Sustainable Vineyard Management",*Communications in Soil Science and Plant Analysis,2013, ISSN: 0010-3624 (Print) 1532-2416*

[12]. Yong Hea, Min Huang, Annia Garc´ıa , Antihus Hernandez , Haiyan Song," Prediction of soil macronutrients content using near-infrared spectroscopy",*ELSEVIER,2007, Computers and Electronics in Agriculture 58 (2007) 144–153*

[13]. Guy Schwartz1, Gil Eshel and Eyal Ben-Dor,"Reflectance Spectroscopy as a Tool for Monitoring Contaminated

[14]. Soils", *Soil Contamination, 2011, ISBN 978-953-307-647-8*

[15]. Junjie Wang , Lijuan Cui ,Wenxiu Gao , Tiezhu Shi , Yiyun Chen, Yin Gao," Prediction of low heavy metal concentrations in agricultural soils using visible and near-infrared reflectance spectroscopy", *ELSEVIER, 2014, Geoderma 216 (2014) 1–9*

[16]. A. Shamsoddini, S. Raval, R. Taplin," Spectroscopic Analysis Of Soil Metal Contamination Around A Derelict Mine Site In The Blue Mountains, Australia", *ISPRS Annals of the Photogrammetric, Remote Sensing and Spatial Information Sciences, 2014, Volume II-7, 2014*

[17]. D. Miranda Salazar, H. L. Mart´ınez Reyes, M. E. Mart´ınez-Rosas, M. M.Miranda Velasco, E. Arroyo Ortega," Visible-near infrared spectroscopy to assess soil contaminated with cobalt",*ELSEVIER,2012, Procedia Engineering 35 (2012) 245 – 253*

[18]. Vaibhav N. Lokhande, Dr. Ratnadeep R. Deshmukh," Monitoring Arsenic Contamination in Agriculture Soil using Spectroradiometer", *IJIRSET, 2016, ISSN (Online): 2319-8753*

[19]. Vaibhav N. Lokhande, Dr. Ratnadeep R. Deshmukh," Identification of Arsenic Content in Soil Using Spectral Signature", *IJIRCCE, 2016, Vol. 4, Issue 6, ISSN (Online): 2320-9801*

[20]. Grzegorz Siebielec, Gregory W. McCarty, Tomasz I. Stuczynski, and James B. Reeves," Near- and Mid-Infrared Diffuse Reflectance Spectroscopy for Measuring Soil Metal Content", *Journal of Environmental Quality, 2004, 33:2056–2069 (2004).*

[21]. Junfeng JiA, Yinxian Song, Xuyin Yuan and Zhongfang Yang," Diffuse reflectance spectroscopy study of heavy metals in agricultural soils of the Changing River Delta, China", *World Congress of Soil Science, Soil Solutions for a Changing World,2010*

[22]. T H O M A S K E M P E R , A N D S T E F A N S O M M E R," Estimate Of Heavy Metal Contamination In Soils After A Mining Accident Using Reflectance Spectroscopy", *Environmental Science & Technology,2002, Vol. 36, No. 12, 2002*

[23]. Silvia Rita Stazi, Francesca Antonucci, Federico Pallottino, Corrado Costa, Rosita Marabottini, Maurizio Petruccioli & Paolo Menesatti," Hyper spectral Visible–Near Infrared Determination of Arsenic Concentration in Soil", *Communications in Soil Science and Plant Analysis,2014, ISSN: 0010-3624 (Print) 1532-2416 (Online)*

[24]. REN Hong-Yan, ZHUANG Da-Fang, A. N. SINGH, PAN Jian-Jun, QIU Dong-Sheng, and SHI Run- Estimation of As and Cu Contamination in Agricultural Soils Around a Mining Area by Reflectance Spectroscopy: A Case Study", *ELSEVIER,2009, Pedosphere 19(6): 719–726, 2009*

[25]. Sonia Gannouni, Noamen Rebai, Saadi Abdeljaoued,"A Spectroscopic Approach to Assess Heavy Metals Contents of the Mine Waste of Jalta and Bougrine in the North of Tunisia", *Journal of Geographic Information System,2012, 4, 242-253*

[26]. R. Casa , F. Castaldi , S. Pascucci , A. Palombo , S. Pignatti," A comparison of sensor resolution and calibration strategies for soil texture estimation from hyper spectral remote sensing",*ELSEVIER,2013,Geoderma 197–198 (2013) 17–26*

[27]. Bo Stenberg," Effects of soil sample pretreatments and standardized rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon",*ELSEVIER,2010, Geoderma 158 (2010) 15–22*

[28]. H.M. Bartholomeus , M.E. Schaepman , L. Kooistra , A. Stevens ,W.B. Hoogmoed , O.S.P. Spaargaren",Spectral reflectance based indices for soil organic carbon quantification,*ELSEVIER,2008, Geoderma 145 (2008) 28–36*

[29]. A. Volkan Bilgili ,Fevzi Akbas , Harold M. van Es," Combined use of hyper spectral VNIR reflectance spectroscopy and kriging to predict soil variables spatially*", Springer, 2011,Precision Agric (2011) 12:395–420*

[30]. HE Ting WANG Jing LIN Zongjian CHENG Ye," Study on Spectral Features of Soil Organic Matter", *The International Archives of the Photogrammetric, Remote Sensing and Spatial Information Sciences, 2008, Vol. XXXVII. Part B7. Beijing*
[31]. Prem Kumar Chandrakar, Sanjay Kumar and Dewashish Mukherjee", Applying Classification Techniques In Data Mining In Agricultural Land Soil", *IJCE, 2011, Volume 3, Number 2, pp. 89–95*
[32]. P. Bhargavi, Dr. S. Jyothi," Soil Classification Using Data Mining Techniques: A Comparative Study", *International Journal of Engineering Trends and Technology, 2011, ISSN: 2231-5381*
[33]. Jay Gholap, Anurag Ingole, Jayesh Gohil, Shailesh Gargade, Vahida Attar," Soil Data Analysis Using Classification Techniques and Soil Attribute Prediction",*2011*
[34]. Jay Gholap," Performance Tuning Of J48 Algorithm For Prediction Of Soil Fertility", *IJCSIT, 2012, ISSN 2249-5126*
[35]. Rashmi Arora, Savneet Kaur," Mining of Soil Data Using Unsupervised Learning Technique", *International Journal of Applied Engineering Research, 2012, ISSN 0973-4562 Vol. 7 No.11 (2012)*
[36]. Vrushali Bhuyar," Comparative Analysis of Classification Techniques on Soil Data to Predict Fertility Rate for Aurangabad District", *IJETTCS, 2014, Volume 3, Issue 2, ISSN 2278-6856*
[37]. Velide Phani kumar and Lakshmi Velide, "Data Mining plays a key role in soil data analysis of Warangal region*", International Journal of Scientific and Research Publications,2014, Volume 4, Issue 3, ISSN 2250-3153*
[38]. M.C.S.Geetha," A Survey on Data Mining Techniques in Agriculture", *IJIRCCE, 2015, Vol. 3, Issue 2, ISSN (Online): 2320-9801*
[39]. Dr. S.Hari Ganesh, Mrs. Jayasudha," Data Mining Technique to Predict the Accuracy of the Soil Fertility", *IJCSMC, 2015, Vol. 4, Issue. 7, pg.330 – 333*
[40]. M.Suriyapraba,S.Sasikala," Different Type of Considerations on Soil Classification using Special Classification Techniques", *International Journal on Recent and Innovation Trends in Computing and Communication,2016, Volume: 4 Issue: 3 ISSN: 2321-8169*