SOIL TEST MINING FOR CULTIVATION FORUM WITH ENHANCED ECLAT APPROACH

Dr. R.B.V.Subrahmanyam[1], T.V.Mahendra[2]
[1]Professor, Dept.of CSE, NIT, Warangal
[2]Associate Professor, Dept.of CSE,NBKRIST,Vidyanagar.

ABSTRACT

Data Mining Techniques have dominated over various methods to gain knowledge from vast amount of data. There are the various research tools available for the large amount of data. Data mining has more technique to analyze the large amount of data like various classification algorithms.We obseve that our national economy highly influenced with agriculture based industries. In this paper ,we propose an improved version of Equivalence Class Transformation(ECLAT)Rule mining technique,which is applied over agriculture land soil test reports to generate advisory details which facilitate decision support to crop rotation, fertilizer requirements and harvesting procedures. The main aim is to avoid unnecessary fertilizer usage during cultivation of lands and to increase soil vitality. With this approach,we can improve crop productivity and provide best nutrients for production.

Key Words: soil-testing, crop-rotation, KDD process, classification,Minerals, organic matter, Mineral ratios, Yield factor.

INTRODUCTION

Data mining has been used to analyze large data sets and establish useful classification and patterns in the data sets. “Agricultural and biological research studies have used various techniques of data analysis including, natural trees, statistical machine learning and other analysis methods. The analysis of agricultural data sets with various data mining techniques may yield outcomes useful to researchers in the Agricultural field. . Data Mining software applications includes various methodologies that have been developed by both commercial and research centers. These techniques have been used for industrial, commercial and scientific purposes. Agricultural and biological research studies have used various techniques of data analysis including natural trees, statistical machine learning and other analysis tools. This research determined whether data mining techniques could also be used to classify soils that analyze large soil profile experimental datasets. The research aimed to establish if data mining techniques can be used to analyze different classification methods by determining whether meaningful pattern exists across various soil profiles characterized at various research sites. Agriculture and allied sectors provide largest livelihood in India. Itcontribution to Gross Domestic Product (GDP)[1] is significant. Supporting rural employment, food security, soil conservation, biodiversity protection, natural resource management activities through agriculture empowers nation economically. Indian government concentrated on activities like*Green Revolution*, *White Revolution*, *Yellow Revolution* and *Blue Revolution* [2,5] to improve agriculture industry. Soil testing is a pre-cultivation activity which gives a brief idea about soil structure and mineral compositions ratios. Surveying this reports agriculture scientists provide guidelines for crop harvesting[6] and fertilizer usage in harvesting. Soil testing [1, 2] must be done frequently to analyze the vitality of soils during the seasonal changes. The essential nutrients required for various crop growths can be estimated during soil testing process. Indian government

providing various free soil testing centers at every district head quarters region wide. In this paper we applied Data Mining techniques over agriculture statistical trained data sets to discover interesting patterns. The main goal of the paper is to assist farmers in crop rotation schemes, fertilizer usage and soil vitality conservation.

2. SOIL PARAMETERS

The nutrients of soil are classified into three categories as shown in Table 1. Mineral ratios[1,4]are also depends on type of soil available in that region. Soil types [3] in entire India are classified into six categories.

Major	Secondary	Minor
Nitrogen(N)	*Sulphur*	*Iron*
Phosphorus(P)	*Calcium*	*Boron*
Potassium(K)	*Magnesium*	*Zinc*
		Manganese
		Copper
		Molybdenum
		Chlorine

Table 1. Plant Nutrients in Soil

The nutrients of soil are classified into three categories as shown in Table 1. Mineral ratios[1,4] are also depend on type of soil available in that region. Soil types[3] in entire India are classified into six categories

Alluvial Soils:

These are formed by the deposition of sediments by rivers. They are rich in humus and very fertile. They are found in Great Northern plain, lower valleys of Narmada and Tapti and Northern Gujarat. These soils are renewed every year.

Black Soils:

These soils are made up of volcanic rocks and lava-flow. It is concentrated over Deccan Lava Tract which includes parts of Maharashtra, Chhattisgarh, Madhya Pradesh, Gujarat, Andhra Pradesh and Tamil Nadu. It consists of Lime, Iron, Magnesium and also Potash but lacks in Phosphorus, Nitrogen and Organic matter[6].

Red Soils:

These are derived from weathering of ancient metamorphic rocks of Deccan Plateau. Its redness is due to iron composition. When iron content is lower it is yellow or brown. They cover almost the whole of Tamil Nadu, Andhra Pradesh, Chhattisgarh, Karnataka, Maharashtra and parts of Orissa.

Laterite Soils:

These soils are formed due to in-tense leaching and are well developed on the summits of hills and uplands. They are commonly found in Kerala, Tamil Nadu, Maharashtra, Chhattisgarh and hilly areas of Orissa and Assam.

Mountain Soils:

These soils are formed as a result of the accumulation of organic matter derived from forest growth. They are found in Himalayan region and vary in different regions according to altitude. Tea is grown in those areas which receive sufficient rainfall.

Desert Soils:

In the desert regions of Rajasthan, soils are not well developed. As evaporation is in excess of rainfall, the soil has a high salt content and saline layer forms a hard crust. These soils are generally sandy and deficient in organic matter.

Testing Process :

Extraction process is preliminary stage for soil testing. Based on land nature to collect soil sample 5cm to 32 cm drilled into ground. For best results random samples must be extracted from lands. The Storage of samples is vital stage in this process. As shown in Fig 1 sealed tubes under 30˚C used to preserve soil nature. Soil samples shouldn't expose to air. Transport stage is crucial where samples are carefully transported to testing labs.

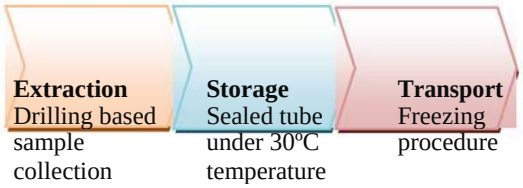


Figure 1. Soil Testing Process

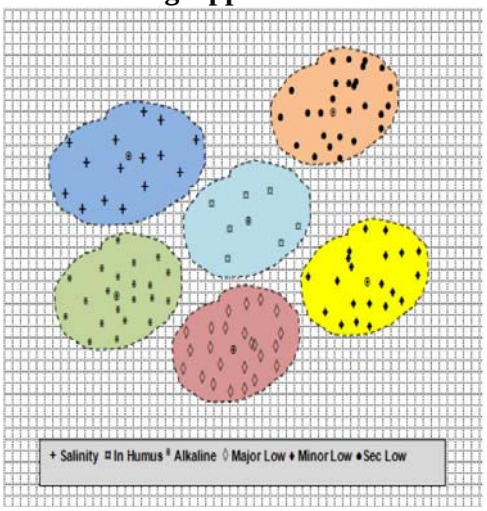
3. COLLECTION OF DATA

In this paper a collection of soil testing reports are pre-processed to generate training data sets. The training data sets are subjected to data mining techniques like ARM, Grouping and clustering. Finally decisions tracked based on group characterization rules implemented. The training data samples Mineral ratios, crop cultivation used to generate various interesting measures for decision support on cultivation.

4. EXPERIMENTAL ANALYSIS

A collection of soil reports from various agriculture lands across south region districts located over Andhra-Pradesh state are collected. A data set of 60 soil testing samples pre-processed to a format acceptable by data mining tool WEKA and TANAGRA. The training sample data file is in .CSV format

4.1 Clustering Application



Cluster centroids

Attribute	Cluster n°1	Cluster n°2	Cluster n°3	Cluster n°4	Cluster n°5	Cluster n°6
N	5.000000	5.000000	5.000000	5.000000	15.000000	15.000000
P	5.000000	5.000000	5.000000	9.000000	20.000000	10.000000
K	5.000000	5.000000	10.000000	6.000000	15.000000	15.000000
Ca	30.000000	5.000000	15.000000	15.000000	25.000000	2.000000
Mg	20.000000	5.000000	10.000000	15.000000	25.000000	3.000000
S	30.000000	5.000000	10.000000	10.000000	20.000000	2.000000
Fe	5.000000	10.000000	10.000000	10.000000	2.000000	10.000000
B	5.000000	5.000000	20.000000	10.000000	4.000000	10.000000
Ma	5.000000	5.000000	25.000000	10.000000	3.000000	10.000000
Zn	5.000000	10.000000	15.000000	20.000000	3.000000	15.000000
Mo	5.000000	10.000000	15.000000	10.000000	3.000000	10.000000
Co	5.000000	10.000000	25.000000	10.000000	3.047619	10.000000
Cl	40.000000	10.000000	20.000000	5.000000	4.857143	5.000000

Computation time : 15 ms.
Created at 1/11/2016 8:22:12 PM

Figure 3. Cluster Analysis of soil tests region based

The application of ECLAT with the depth first search mechanism to soil test reports collected over southern region distict zones resulted in formation of six major clusters based on soil nutrients ratios as shown in Fig. 3 above. Above analysis gives report that 15% of salinity, 6% In-Humus, 21% Alkaline, 25% Major-low, 21% Minor-low and 29% Secondary-low lands are distributed over regional districts.

4.2 IMPROVED ECLAT Approach

The application of enhanced Equivalence class Transformation rule mining technique on crops training set data resulted interesting rules with best crop nutrients requirement. The yield of crop also increased by suggesting best crop rotation for farmer along with estimated rotation schemes (Frequent Item sets) to assist farmers to improve farm productivity. Apriori algorithm applied on previous crop rotation data on the basis of yield factor [6].

Results

ITEMS

Transactions	24
Counting items	
All items	6
Filtered items	6
Counting itemsets	
card(itemset) = 2	12
card(itemset) = 3	10
card(itemset) = 4	3
Rules	
Number of rules	**9**

RULES

Number of rules : 9

N°	Antecedent	Consequent	Lift	Support (%)	Confidence (%)
1	"Legumes=true" - "Umbellifers=true"	"Hcrop=true" - "Alliums=true"	2.00000	20.833	100.000
2	"Legumes=true" - "Hcrop=true" - "Umbellifers=true"	"Alliums=true"	2.00000	20.833	100.000
3	"Umbellifers=true"	"Legumes=true" - "Hcrop=true" - "Alliums=true"	2.00000	20.833	100.000
4	"Hcrop=true" - "Umbellifers=true"	"Legumes=true" - "Alliums=true"	2.00000	20.833	100.000
5	"Legumes=true" - "Umbellifers=true"	"Alliums=true"	2.00000	20.833	100.000
6	"Umbellifers=true"	"Legumes=true" - "Alliums=true"	2.00000	20.833	100.000
7	"Hcrop=true" - "Umbellifers=true"	"Alliums=true"	2.00000	20.833	100.000
8	"Umbellifers=true"	"Hcrop=true" - "Alliums=true"	2.00000	20.833	100.000
9	"Umbellifers=true"	"Alliums=true"	2.00000	20.833	100.000

Computation time : 0 ms.
Created at 2/22/2016 10:18:56 PM

Figure 4. Association Rule Mining on Crop Rotation Data

Using these rules we can find best fit crop rotation for fields in various locations, also we can estimate the total annual plant nutrients requirement. The yield of crop also increased by suggesting best crop rotation for farmer along with estimated fertilizer usage reports. For example to harvest heavy crops farmer should use {Legumes, Umbellifers, Alliums} crops compulsory to ensure good yield of crops. To increase the soil nutrition {Alliums, Umbellifers} crop rotation best suited. For both Kharif and Rabi seasons this process can generate rules for best crop rotation.

4.3 Equivalence Classes for Soils

Group characterization 1

Parameters

Results

Description of "class"

class=Salinity

Examples		[12.8 %] 15	
Att - Desc	Test value	Group	Overral
Continuous attributes : Mean (StdDev)			
N	-3.55	5.00 (0.00)	9.27 (4.97)
P	-3.78	5.00 (0.00)	9.79 (5.23)
K	-5.16	5.00 (0.00)	10.38 (4.31)
Discrete attributes : [Recall] Accuracy			

class=Inhumus

Examples		[5.1 %] 6	
Att - Desc	Test value	Group	Overral
Continuous attributes : Mean (StdDev)			
N	-2.15	5.00 (0.00)	9.27 (4.97)
P	-2.29	5.00 (0.00)	9.79 (5.23)
K	-3.13	5.00 (0.00)	10.38 (4.31)
Discrete attributes : [Recall] Accuracy			

class=Alkaline

Examples		[17.9 %] 21	
Att - Desc	Test value	Group	Overral
Continuous attributes : Mean (StdDev)			
K	-0.45	10.00 (0.00)	10.38 (4.31)
N	-4.33	5.00 (0.00)	9.27 (4.97)
P	-4.61	5.00 (0.00)	9.79 (5.23)
Discrete attributes : [Recall] Accuracy			

class=Major-low

Examples		[21.4 %] 25	
Att - Desc	Test value	Group	Overral
Continuous attributes : Mean (StdDev)			
P	-0.84	9.00 (0.00)	9.79 (5.23)
N	-4.83	5.00 (0.00)	9.27 (4.97)
K	-5.71	6.00 (0.00)	10.38 (4.31)
Discrete attributes : [Recall] Accuracy			

class=Sec-low

Examples		[24.8 %] 29	
Att - Desc	Test value	Group	Overral
Continuous attributes : Mean (StdDev)			
N	7.13	15.00 (0.00)	9.27 (4.97)
K	6.62	15.00 (0.00)	10.38 (4.31)
P	0.25	10.00 (0.00)	9.79 (5.23)
Discrete attributes : [Recall] Accuracy			

Computation time : 0 ms.
Created at 2/23/2016 1:31:42 AM

Figure 5. Group Discrimination of Agriculture Lands

In this experiment soil test reports collected over southern regions of Andhra Pradesh. The major nutrients for plant growth are N, P and K. The data mining tool performed discrimination of agriculture lands into five major classes namely Salinity, In-humus, Major-low and Minor-low as shown in Figure 5.The statistical analysis assists to get effective crop rotation mechanism based on the nutrient values.

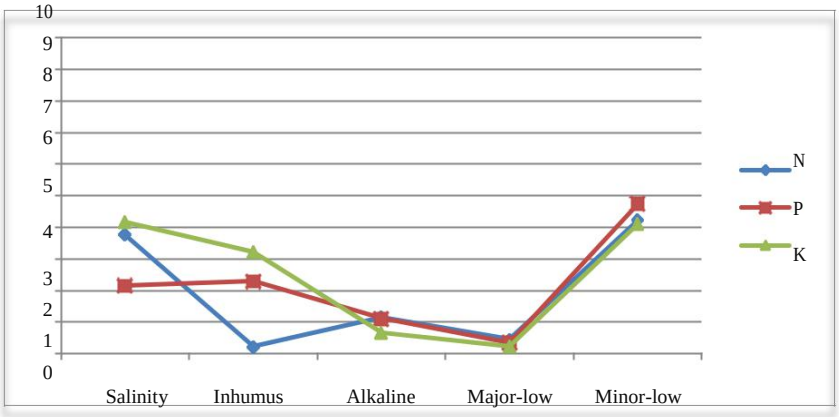


Figure 6. Ratio of minerals over lands (0.1=1 Sq. m/Hector)

From the Figure 6 graph we can depict the ratio of NPK in soil over various groups of soils collected over the cultivation lands of AP and Telangana. Using this statistics dynamically Agriculture server can send fertilizer usage suggestions to respective mandals in districts. Also we can provide optimal crop rotation schemes to respective class of lands. The NPK values must cross 5 for a healthy cultivation land. Farmers can depend on fertilizers or Bio-Fertilizers (Eco-Friendly) to improve the ratio of NPK as per suggestions provided by agriculture department through web site.

5. CONCLUSION AND FUTURE SCOPE

The experiments conducted in this paper concluded that using Data Mining tools in the analysis of Agriculture data such as Soil Test reports improves the decision making quality of agriculture officers. Integration of data mining techniques with Agriculture Server turns it into an automated cultivation advisor and reduces laborious manual statistic analysis. Also improves accuracy in information extraction process. In future we are going to expand the crop rotation at various locations across notation for healthy environment and also to build automated software tool based framework.

REFERENCES

[1] Jay Gholap, Anurag Ingole, Jayesh Gohil, Shailesh Gargade and Vahida Attar, "*Soil Data Analysis Using Classification Techniques and Soil Attribute Prediction*", IJCSI, Vol. 9, Issue 3, No 3, ISSN: 1694-0814, May 2012.

[2] Jay Gholap, Anurag Ingole, Jayesh Gohil, Shailesh Gargade, Vahida Attar, "*Soil Data Analysis Using Classification Techniques and Soil Attribute Prediction*", Press-College of Engineering, Pune.

[3] Dr. S.Hari Ganesh and Mrs. Jayasudha, "*Data Mining Technique to Predict the Accuracy of the Soil Fertility*", IJCSMC, Vol. 4, Issue. 7, July 2015, pg.330 – 333.

[4] "*Interpretation of Soil Test Reports for Agriculture*", www.msuextension.org/store, Revised October -2013, 1013SA.

[5] Swanti.A.Jain, M.S.Jagtap and K.P.Patel, "*Physico-Chemical Characterization of farmland Soil used in some villages of Lunawada Taluka. Dist: Mahisagar (Gujarat) India*", International Journal of Scientific and Research Publications, Volume 4, Issue 3, March 2014, ISSN 2250-3153.

[6] JayGholap, "*Performance Tuning of J48 Algorithm for Prediction of Soil Fertility*", Asian Journal of Computer Science And Information Technology, ISSN: 2249-5126, pp: 251-252, 2012.