Soil Classification Support for Farming Decisions

Prof. Nilesh Patil
Department of Information Technology
Rajiv Gandhi Institute of Technology
Mumbai, India
e-mail: nileshdeep@gmail.com

Apoorav Singh
Department of Information Technology
Rajiv Gandhi Institute of Technology
Mumbai, India
e-mail:apooravslakra@gmail.com

Piyush Patel
Department of Information Technology
Rajiv Gandhi Institute of Technology
Mumbai, India
e-mail:patelpiyush567@gmail.com

Nisarg Joshi
Department of Information Technology
Rajiv Gandhi Institute of Technology
Mumbai, India
e-mail:nisargj17@gmail.com

Pranav Pawaskar
Department of Information Technology
Rajiv Gandhi Institute of Technology
Mumbai, India
e-mail:pranavpawaskar@gmail.com

Abstract**—**The techniques of data mining are extremely popular in the area of agriculture. Data mining involves the systematic analysis of huge information sets, and data processing in agricultural soil datasets is exciting and fashionable analysis space. The productive capability of a soil depends on soil fertility. Today, data processing is employed in a very large area and plenty of ready-to-wear data processing system product and domain specific data processing application software's are obtainable, however data processing in agricultural soil knowledge sets may be a comparatively a young analysis field. In this paper, we offer internet base answer for the soil testing laboratories yet as free messages for the farmer that contains data like soil testing code, chemical that is important for the crop and additionally the knowledgeable recommendation. Additionally, farmers specify their next crop whereas they furnish their sample to scantiest therefore in keeping with next crop the chemical can recommend. The results supported the classification of contains that should be gif tin soil and in keeping with result report are generated.

*****

I. Introduction

Data Mining is a very crucial research domain in recent research world. The techniques are useful to elicit significant and utilizable knowledge which can be perceived by many individuals. Data Mining Software application includes various methodologies that have been developed by both commercial and research center. These techniques have been used for industrial, commercial and scientific purposes. Agricultural and biological research studies have been used for various techniques of data analysis including, natural trees, statistical machine learning and other analysis methods. This research aimed to assess data mining techniques and apply them to a soil science database to establish if meaningful relationships can be found.

Efficient techniques can be developed and tailored for solving complex soil data sets using data mining to improve the effectiveness and accuracy of the Classification of large soil data sets. A soil test is the analysis of a soil sample to determine nutrient content, composition and other characteristics. Tests are usually performed to measure fertility and indicate deficiencies that need to be remedied. The soil testing laboratories are provided with suitable technical literature on various aspects of soil testing, including testing methods and formulations of fertilizer recommendations. It helps farmers to decide the extent of fertilizer and farm yard manure to be applied at various stages of the growth cycle of the crop. Soil can be characterized using physical properties: EC, Ph, bulk density etc. and chemical properties that can be sub divided into macro nutrients: phosphorous, potassium and nitrogen, micro nutrients: zinc, iron, copper etc. Various classification algorithms like ID3 are employed to classify the soil and indicate the essential requirements for the soil.

II. Objectives

The objectives here is to implement a robust system designed in such a way that it classifies soilbased on its chemical components for farmer using modern technical expertise. Also, we aim tominimize the efforts made to maintain and update the entire system. By using latest datasets anda classification algorithm which suits our purpose we wish to design a system which can guide afarmer and aid his decisions in the area related to soil and crops.

III. System Description

Soil Classification Support will mainly be helpful for farmers who want soil to be analyzed. This will be done by taking into consideration a few indicators such as pH, Nitrogen, Phosphorous, and Potassium levels. This analysis will help make crucial farming decisions. The proposed system is very different from the existing system. Currently there is no application support to help farmers with soil analysis. The proposed system will be fast, scalable, more reliable and dynamic. The proposed system will consist of windows application software. We will upload a dataset in the software and train the system. After training, the system will generate a set of rules regarding the ID3 algorithm. The set of rules generated will be saved in the database. After that the user can enter the data. Based on the set of rules generated, the data will be analyzed by the system. After analyzing the data, the system will give the output and provide suggestions to the user regarding the production of crop in his field. Given below is the basic diagram for the proposed system.

IV. LITERATURE REVIEW

Proper management of all essential nutrients is critical both in developed and developing nations. Better management of nutrients and field-specific recommendations will lead to improved food production while limiting environmental degradation (Goulding et al., 2008). However, having numerous competing uses, the presence of biomass and soil organic matter (SOM) enhance benefits of any inorganic fertilizers added to the field. Inorganic and organic fertilizers when implemented, can add valuable inputs to the soil. When coupled with soil conservation practices, the fertilizers can be retained in the topsoil and provide nutrients to the subsequent crops (Fresco and Jager, 2013).

Nutrient loss due to erosion transport is a major source of declining soil fertility in SSA. In 1996, the total loss of N and P were 109 and 13 kg ha-1 respectively (Singh and Lal, 2005). Since that time, total N, P, K losses increased in SSA to 8 million Mg in 2004 (Henao and Baanante, 2006). Nutritional constraints that lead to low productivity are primarily: low levels of N, P, and K, combined with low cation exchange capacity (CEC) (Kaihura et al., 1999). Additional factors that cause decreased productivity are: reduced plant available water, degraded soil structure, naturally low fertility, and a loss of SOM and lack of plant available N (NRC, 2010). Total Nitrogen loss caused by soil erosion was 50-100 times the loss from runoff in a study done in 1996 (Singh and Lal, 2005).

Application of farm yard manure (FYM) can increase soil pH, SOC, and plant available nutrients, and in turn maize yield, in a study of Tanzanian soil (Kaihura et al., 1999). Unlike inorganic fertilizer, FYM can be locally sourced, and is much less expensive. This study found that the application of FYM could reduce soil erosion and improve soil quality and aggregation (Kaihura et al., 1999). Nitrogen in the soil is made plant-available through soil water moving through pores and diffusing quickly as nitrate-N. Soil structure and aggregates protect the SOM from decomposing quickly. The SOM when decomposed slowly, mineralizes nutrients to bioavailable forms. The texture of the soil also determines the rate of SOM decomposition (Goulding et al., 2008). Highly compacted or crusted soils can cripple nitrogen by the anaerobic conditions and denitrification. SOM contains and protects most of the nitrogen and nearly all the phosphorous and sulfur that is supplied to crops (Goulding et al., 2008). Nutrient content of Tanzanian soils is inversely related to the rate of soil erosion. FYM is generally the finest soil input, its benefits often exceeding the application of N or P fertilizers (Kaihura et al., 1999).

Organic resources and mineral fertilizers, when used together, offer improved crop yields, soil fertility and can aggregate soil management practices. One reason for this improvement in yield could be accredited to the organic resources halting the mineral fertilizer until the plant is ready for consumption. Soils with extremely low organic carbon content (below the 1.5% threshold level) may respond poorly to the addition of mineral fertilizers and other inputs (Lal 2010). Organic resources that have a high N level, like compost and manure can increase maize yields more than other inputs, specifically, the addition of N fertilizer, maize stover, the control, or sawdust (Chivenge et al., 2007).

Manure addition brought about a doubling of maize yield when the manure application rate is increased from 5 to 10 Mg ha-1 (Kimani et al., 2007). The nutrients lost to plant uptake or erosion are balanced by such inputs, shrinking the yield gap, regardless of the crop grown.

Greater agronomic productivity is achieved by solving soil degradation problems, and conservation targeted at the smallholder farms eliminates some of the most critical degradation and nutrient depletion present (Sanchez and Swaminathan, 2005). The major cause of low food production in SSA is because of decreasing soil fertility (Sanchez, 2002). The smallholder farming systems compromising 80% of all farms in SSA, when defined as smaller than 2 hectares in size. There are 33 million smallholder farms in SSA that could hypothetically adopt new agriculture (Wiggins, 2009). The viable impact of slowing soil nutrient mining and soil erosion on smallholder systems could have an enormous impact on food security in the province.

The Kilimanjaro Region of Tanzania is comprised of many different soil types and the parent materials vary from alluvium ash fall from Mt. Kilimanjaro’s eruption, to ancient lake beds, to present-day floodplains.It is defined by monsoonal rains, the bimodal rainfall pattern, amount of soil moisture present, and the tropical semi-arid topographical classification (USDA, 2010).

Few published studies evaluate the effects of permanent cropping on the soils of Tanzania. However, Hartemink (1996) evaluated recent (1990) soil in Tanga, of northern Tanzania, and compared the historical soil analysis of samples evaluated for the same parameters in the 1950s and 1960s. Acrisols were one of the soil groupings studied and the pH had

 ISSN: 2321-8169
221 – 225

dropped one pH unit in 25 years of sisal (*Agave sisalana L.*) production. Acrisols were one of the soil groupings most altered by crop production in this study, showing a low resilience. A different study evaluating the effects of deforestation of Tanzania and the impact on soil fertility, showed a decrease of organic carbon content by 50% when the forest was cleared. This was the most drastic reduction of organic carbon studied, other deforestation locations with younger parent materials, showed a much smaller response. Bulk density and CEC values for the ultisol also indicated the largest response, bulk density increasing and CEC lowering more than all other systems evaluated (Allen and Mar, 1985).
Rainfall is bimodal and recently has been variable with most of the annual precipitation occurring during the long rains from that begin in February and end in May or June. This long rainy season is known in Kiswahili as *masika*. This time is considered the principal and most important growing season of the year. A second growing season occurs from July to December which depends solely on irrigation with a supplemental amount of rainfall. *Vuli*, as this season is known to locals, is less predictable and rain intensity and duration vary year to year. The short rains are the period where farmers plant and grow maize, however crop failure is common, and many maize plants have cobs with unfilled or empty grains. Aggregation of weather parameters from a local meteorological station confirm Lower Moshi Irrigation Scheme has an average annual rainfall of 485 mm and the daily evaporation rate averages 5 mm.
The original USLE was developed in the United States to demonstrate and validate soil conservation measures (Wischmeier and Smith, 1978). It considers inputs of remote sensing data, soil texture information, spatial rainfall patterns, slope and slope length. It is the most widely used soil loss model and has been revised many times since its development. Two successive models are the Revised Soil Loss Equation developed by Renard (1997) and Modified Soil Loss Equation by Williams (1977) and use high resolution inputs.
Kabanza (2013) found soil loss of inland plains of Tanzania to range from 3.7 Mg ha-1 to 11.6 Mg ha-1. One calculation of the USLE in Northern Tanzania agricultural land resulted in the soil loss range of 2.01 Mg ha-1 to 12.34 Mg ha-1 (Ndomba, 2010). Other soil loss models can be found in the literature; however, most are concentrated in highland and mountainous regions of Tanzania.

V. IMPLEMENTATION

The system is in form of desktop based application which helps in farming decision on basis of soil fertility level and gives crop suggestions. Using this system, the farmer can create an account and then login into the system. After login farmer first need to enter the parameters into the various fields of the system. Then this entered parameters are compared to the database of the system. And system gives the soil fertility level and crop suggestions using implementing ID3 algorithm.
Soil classification system is essential for the identification of soil properties. Expert system can be a very powerful tool in identifying soils quickly and accurately. Traditional classification systems include use of tables, flow-charts. This type of manual approach takes a lot of time, hence quick, reliable automated system for soil classification is needed to make better utilization of technician's time.
We propose an automated system that has been developed for classifying soils based on fertility. Being rule-based system, it depends on facts, concepts, theories which are required for the implementation of this system. Rules for soil classification were collected from soil testing lab. The soil sample instances were classified into the fertility class labels as: High, Medium, Low. These class of labels for soil samples were obtained with the help of this system and they have been used further for comparative study of classification algorithm.
The classification of soil was considered critical study because depending upon fertility class of the soil domain knowledge experts determines which crop should be taken on that particular soil.
ID3 is a simple decision tree learning algorithm developed by Ross Quinlan (1983). The basic idea of ID3 algorithm is to construct the decision tree by employing a top-down, greedy search through the given sets to test each attribute at every tree node. In order to select the attribute that is most useful for classifying a given sets, we introduce a metric-information gain.
To find an optimal way to classify a learning set, what we need to do is to minimize the questions asked (i.e. minimizing the depth of the tree). Thus, we need some function which can measure which questions provide the most balanced splitting.

Figure 1 Login page

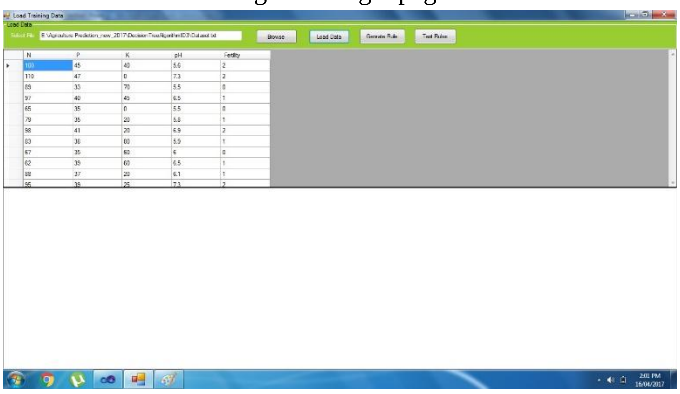

Figure 2 Dataset is loaded.

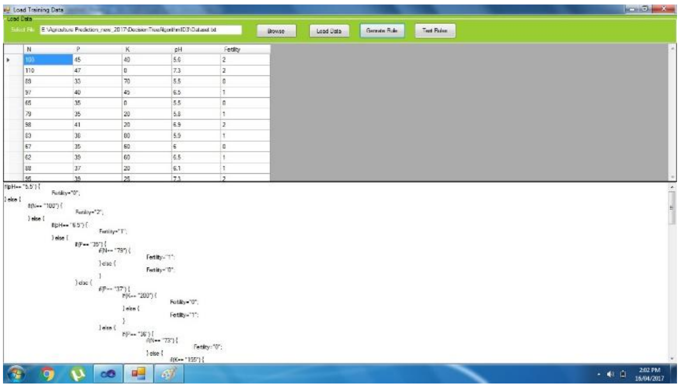

Figure 3 System is trained.

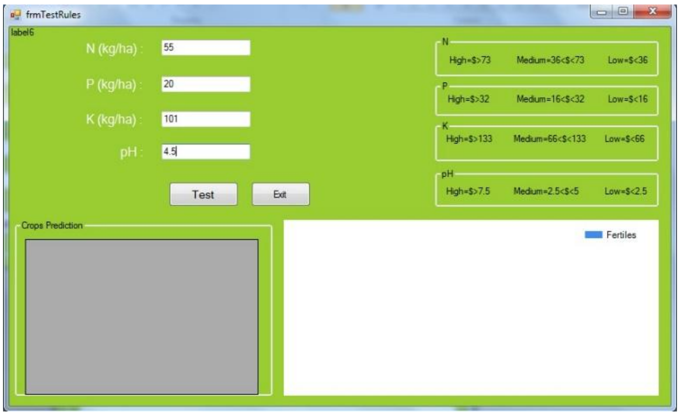

Figure 4 Parameters are entered in the system.

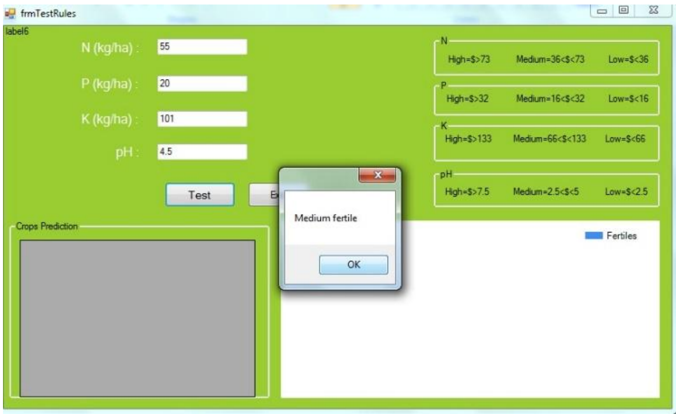

Figure 5 Output is displayed.

Figure 6 Crops are suggested based on fertility level.

VI. Conclusion and Future Scope

In this system, we have proposed an analysis of the soil data using ID3 algorithm as a simple classifier to make a decision tree. The proposed system will help farmers decide which crops are most suitable for their soil type based on classification. The proposed system will help us replace the current system which is quite tiresome and difficult to use. By using this system, we can actually manage to make low cost decision support system. In future, we contrive to build Fertilizer Recommendation System which can be utilized effectively by the Soil Testing Laboratories. This System will recommend appropriate fertilizer for the given soil sample and cropping pattern.

References

[1] Ahmed, A.U., R.V. Hill, L.C. Smith, and D.M. Wiesmann. 2007. The World's Most Deprived: Characteristics and Causes of Extreme Poverty and Hunger.

[2] Allen, J.C., and N. Mar. 1985. Soil Response to Forest Clearing in the United States and the Tropics: Geological and Biological Factors Soil Response to Forest Clearing in the United States and the Tropics: Geological and Biological Factors1. 17(1): 15–27.

[3] Andersson, J. a., and S. D'Souza. 2014. From adoption claims to understanding farmers and contexts: A literature review of Conservation Agriculture (CA) adoption among smallholder farmers in southern Africa. Agric. Ecosyst. Environ. 187: 116–132http://linkinghub.elsevier.com/retrieve/pii/S0167880913002685 (verified 28 May 2014).

[4] Anthofer, J. 2004. The Potential of the SRI for Poverty Reduction in Cambodia. Consultancy Report. FSNPSP

[5] Asfaw, S., B. Shiferaw, F. Simtowe, and L. Lipper. 2012. Impact of modern agricultural technologies on smallholder welfare: Evidence from Tanzania and Ethiopia. Food Policy 37(3): 283–295 http://linkinghub.elsevier.com/retrieve/pii/S0306919212000176 (verified 4 June 2014).

[6] Baker, C.J., K.E. Saxton, W.R. Ritchie, W.C.T. Chamen, D.C. Reicosky, and F.R.S.E. Justice. 2006. No-tillage seeding in conservation agriculture (CJ Baker, KE Saxton, WR Ritchie, WCT Chamen, MFS Ribeiro, SE Justice, and PR Hobbs, Eds.). http://www.cabi.org/CABeBooks/default.aspx?site=107&page=45&LoadModule=PDFHier&BookID=486.

[7] Barrett, C.B., C.M. Moser, and O. V Mchugh. 2004. Better Technology, Better Plots or Better Farmers? Identifying Changes in Productivity and Risk Among Malagasy Rice Farmers Better Technology, Better Plots or Better Farmers? Identifying Changes in Productivity and Risk Among Malagasy Rice Farmers. (687): 0–37.

[8] Baudron, F. 2007. Conservation agriculture in Zambia: a case study of Southern Province.

[9] Baudron, F., M. Jaleta, O. Okitoi, and A. Tegegn. 2014. Conservation agriculture in African mixed crop-livestock systems: Expanding the niche. Agric. Ecosyst. Environ.

187: 171–182 http://linkinghub.elsevier.com/retrieve/pii/S0167880913002806 (verified 27 May 2014).

[10] Bindraban, P.S., M. van der Velde, L. Ye, M. van den Berg, S. Materechera, D.I. Kiba, L. Tamene, K.V. Ragnarsdóttir, R. Jongschaap, M. Hoogmoed, W. Hoogmoed, C. van Beek, and G. van Lynden. 2012. Assessing the impact of soil degradation on food production. Curr. Opin. Environ. Sustain. 4(5): 478–488 http://linkinghub.elsevier.com/retrieve/pii/S1877343512001224 (verified 10 July 2014).

[11] Cassman, K.G. 1999. Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture. Proc. Natl. Acad. Sci. U. S. A. 96(11): 5952–9 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=34211&tool=pmcentrez&rendertype=abstract.

[12] Chapagain, a. K., and a. Y. Hoekstra. 2010. The green, blue and grey water footprint of rice from both a production and consumption perspective. Ecological Economics (40): 1–62.

[13] Chauhan, B.S., and D.E. Johnson. 2009. Ecological studies on Cyperus difformis, Cyperus iria and Fimbristylis miliacea: three troublesome annual sedge weeds of rice. Ann. Appl. Biol. 155(1): 103–112http://doi.wiley.com/10.1111/j.1744-7348.2009.00325.x (verified 25 February 2015).

[14] Chivenge, P.P., H.K. Murwira, K.E. Giller, P. Mapfumo, and J. Six. 2007a. Long-term impact of reduced tillage and residue management on soil carbon stabilization: Implications for conservation agriculture on contrasting soils. 94: 328–337.

[15] Chivenge, P., H. Murwira, K. Giller, P. Mapfumo, and J. Six. 2007b. Long-term impact of reduced tillage and residue management on soil carbon stabilization: Implications for conservation agriculture on contrasting soils. Soil Tillage Res. 94(2):328337http://linkinghub.elsevier.com/retrieve/pii/S0167198706001917 (verified 24 May 2014).

[16] Chivenge, P., H. Murwira, K. Giller, P. Mapfumo, and J. Six. 2007c. Long-term impact of reduced tillage and residue management on soil carbon stabilization: Implications for conservation agriculture on contrasting soils. Soil Tillage Res. 94(2):328337http://linkinghub.elsevier.com/retrieve/pii/S0167198706001917 (verified 7 January 2015).

[17] Diao, X., D. Headey, and M. Johnson. 2008. Toward a green revolution in Africa: what would it achieve, and what would it require?Agric.Econ.39:539550http://doi.wiley.com/10.1111/j.1574-0862.2008.00358.x(verified-11-July-2014)