J48 Classifier Approach to Detect Characteristic of Bt Cotton base on Soil Micro Nutrient

Youvrajsinh Chauhan[#1], Jignesh Vania[*2]

[#1]*Department of Computer Engineering, LJIET*
Ahmedabad, India

[*2]*Aassistant Professor of L J Institute of Engineering and Technology*
Ahmedabad, India

***Abstract*—Agriculture is an emerging research field that is experiencing a constant development. In agriculture area problem of efficient knowledge exploitation and acquisition is very popular. In agriculture decision making process is play an important role. The decision tree is one of the common modelling methods to classify. In this paper, we present the basic knowledge of soil nutrients, how affect and related with crop. In this research, Steps for building a predictive model of crop health that have been explained. A large data set of soil nutrient database of Bt Cotoon is extracted from the Soil Micronutrient & Agriculture department, Ananad. The outcome of this research could improve crop production and identify crop disease that help soil systems, uses throughout a large number of fields or area that include environment, agriculture, horticulture and land use management.**

***Keywords*— prediction, soil nutrients, agriculture, decision tree.**

I. INTRODUCTION

Soil fertility, the health of soil for growing plants, is recognized as a primary constraint to agricultural production in developing countries [1]. One parameter to evaluate the sustainability of production systems is soil quality, which includes organic matter, acidity and/or alkalinity, availability and balance of nutrients, structure and aggregation, water infiltration and storage, as well as plant productivity [2]. Under some circumstances, soil fertility is the indicator of soil quality because these variables provide simplified representations of soil fertility like the availability, distribution and balance of nutrients in soil [3]. Soil fertility evaluation and crop health, through the classification of soil nutrients including macro and trace elements, is necessary for sustainable management of land resources.

Decision Tree is the classification technology that generated from disorder and random data, the classification accuracy of which is high and the result pattern is simple. Used various techniques of data analysis including, natural trees, statistical machine learning and other analysis methods that have studied by agricultural and biological research. *Cunningham and Holmes*, 1999. The agriculture analysis of its data sets with various decision tree techniques may yield outcomes useful in the Agriculture field's researchers.

This paper covers the most recent application of decision tree that use on agriculture. This paper is not meant to be exhaustive. We will give particular attention recent works and then we will briefly mention some other application that looked to us to be the most interesting to report and help to do more on forward which are make a path that how to conclude the concepts of J48 decision tree techniques that leads to make true decision making on agriculture. J48 is give higher accuracy result of 91.90 percent for predicts the crop health model other then NBTree and SimpleCART algorithms [13]

This paper is organized as follows. In Section II, we will try to present the concept and overview of nutrients in agriculture that helps to know why nutrient is required for crop and their impact, deficiency, symptoms on crops and give the basic knowledge and how it related with other factor and take overview of J48 decision tree . In Section III, explain decision tree process of J48 tree and give result of it using yellowing crop disease of Bt Cotton. Section IV, give result taken by WEKA J48 tree Some final remarks will be given in Section 5.

II. MATERIAL AND APPROACH

Agricultural field is very wide area. Here some concept leads as theoretical for more understanding and better grasping

A. Nutrients

In Agriculture, proper crop development sixteen nutrients are essential which absorb from soil. Each one is equally important to the plant which required in vastly different amounts. These essential elements differences have led to the grouping into three categories like primary (macro) nutrients, secondary nutrients and micronutrients which all are needed to plant or crop for develop.

B. Nutrient Functions

According [6], Nitrogen is a major component of proteins, vitamins, hormones, chlorophyll and enzymes essential for plant life. Phosphorus is necessary for protein formation, photosynthesis, seed germination and almost all aspects of growth in plants. Potassium is necessary for formation of starches, carbohydrates, protein synthesis, sugars and cell division in roots of the plant. Sulfur is a structural and modelled component of proteins, vitamins, amino acids and enzymes and also it is essential to produce chlorophyll. Critical structural component of the chlorophyll molecule is Magnesium and is necessary for plant enzymes functioning to produce sugars, fats and carbohydrates. Activates enzymes Calcium is a basic structural component of cell walls, influences water movement into cell and necessary for growth and division of cell. Many enzymes functions and the synthesis of chlorophyll is caused and related to Iron. Manganese is involved in respiration, nitrogen metabolism and enzyme activity for photosynthesis. Boron is necessary for cell wall formation, calcium uptake, membrane integrity and aid in the translocation of sugars may be by it. Boron affects 16 functions in plants. These functions include pollen germination, fruiting, flowering, water relationships, cell division and the movement of hormones. Zinc is a component of enzymes of a large number of enzymes including plant growth hormones (auxins). Copper is plays a part in nitrogen metabolism and also concentrated in roots of plants. Molybdenum is reduces nitrates to ammonia that a structural component of the enzyme. Chlorine is involved in movement of water in cells and solutes in cell, the ionic balance for crop or plants to take up mineral elements. Include deficiency symptoms as stubby roots, chlorosis (yellowing), bronzing and wilting. Nickel has just recently won the status as an essential trace element for plants according to the Agricultural Research Service Plant, Soil and Nutrition Laboratory, Ithaca, NY and it is required for the enzyme urease and then liberate the nitrogen. Sodium is involved in ionic balance in plants and osmotic (water movement). Cobalt is required for nitrogen fixation which is in legumes. Silicon is use for cell which is found as a structural component of cell walls. With supplies of soluble silicon in plants produce stronger, tougher cell walls which making them a mechanical barrier and wall to piercing and sucking insects.

C. Nutrient Functions

These two tables namely macro nutrient and micro nutrient deficiency gives the ideas to effect of sixteen nutrients on crop by its availability [19, 20, 21,22].

Nutrient	Deficiency	Excess
Nitrogen (N)	Reduced growth, Light green to yellow foliage. Reds and purples Sometimes with pink tints	Succulent growth, Thick and brittle; leaves are dark green, Poor fruit set; excess ammonia can induce calcium deficiency
Phosphorus (P)	Reduced growth; leaves dark green; purple or red in older leaves,	Shows up as micronutrient deficiency of Fe, Zn, or Co
Potassium (K)	Reduced growth; Shortened internodes; Older leaves become burn; Poor flowering or fruiting	Causes N deficiency in plant and may affect the uptake of other positive ions such as Mg and Ca
Magnesium (Mg)	Yellowish, bronze, Reddish color,	small necrotic spots in older leaves; Interferes with Ca uptake;
Calcium (Ca)	Inhibition of bud growth; roots can turn black and rot; young leaves are scalloped structure is weak;	Interferes with Mg absorption; high Ca usually causes high pH which then precipitates many of the micronutrient so they become unavailable to the plant
Sulfur (S)	Rarely deficient; general yellowing of the young leaves	Sulfur excess is usually in the form of air pollution

Table-2 Macro nutrient with deficiency and excess

Nutrient	Deficiency	Excess /Comments
Iron (Fe)	Interveinal chlorosis Ssoil high in Ca, Poorly drained soil, soil high in Mn, high pH, high P, soil high in heavy metals (Cu, Zn), Dryness of branches and leaves, Yellowing of leaves.	Rare except on flooded soils
Boron (B)	Failure to set seed; Internal breakdown of fruit or vegetable; Death of apical buds, Giving rise to witches broom; Corkiness & Dryness of fruits, Internal rottening in plant	Tips and edges of leaves exhibit necrotic spots coalescing into a marginal scorch
Zinc (Zn)	Young leaves are very small, Small size of leaves, Deformed leaves	Poor germination; entire leaf is affected by chlorosis
Copper (Cu)	New growth small, Misshapen, wilted; Gummosis in plants, Dieback in plant, Discolour of leaf	Can occur at low pH; shows up as Fe deficiency
Mangane	Poor bloom size and color;	Reduction in

se (Mn)	induced by excessively high pH. Discoloration of leaves, Yellowing of leaves	growth, brown spotting on leaves; shows up as Fe deficiency;
Molybdenum (Mo)	Interveinal chlorosis on older or midstem leaves; Twisted leaves (whiptail); Yellow spots in the leaves, Loss of greenness in leaves	Intense yellow or purple color in leaves; rarely observed
Chlorine (Cl)	Yellowing of leaves and fall from plants Fading of plant	Salt injury, leaf burn, may increase succulence
Cobalt (Co)	This need by plants recently established; essential for Nitrogen fixation	Little is known about its deficiency or toxicity symptoms
Nickel (Ni)	Eessential for seed development	

Table -3 Micro Nutrient with deficiency

D. Soil PH

Alkalinity or acidity of a growing media water solution is referred to as PH [6].The pH of the growth medium has significant effects on the properties of soils and consequently on the nutrient uptake by crop plants. Thomas [7] noted that three soil pH ranges are particularly informative: a pH less than 4 indicates the presence of free acids generally from oxidation of sulfides; a pH less than 5.5 suggests the occurrence of exchangeable Al(Alkeline); and a pH from 7.8 to 8.2 indicates the presence of calcium carbonate, an important agent of calcareous soil. Soils with pH values ranging from 4 to 7 [8] are extensively distributed throughout the tropical and subtropical regions of the world. Soil pH is an important factor influencing the growth of most crops and pastures and the distribution of native plant species [9, 10]. Among the various plant parts, the roots growth mediums are directly affected by the pH. A balance of hydrogen to hydroxyl ions yields (=7.0), a pH neutral soil. The range for mostly crops is 5.5 to 6.2 or slightly acidic creates the greatest average level which shows for availability for all essential plant nutrients. Extreme fluctuations of pH at higher or lower level can cause deficiency or toxicity of nutrients [6].

E. Availability of Nutrients

All nutrients are measured throughout soil testing in soil testing laboratories. Soil testing laboratories get result of measurements or rating by different chemical techniques. By taken example of cotton, require nutrients N, P, K for growing of cotton plant is 240-280, 28-56 and 140-180 in (kg ha-1) [11].The all nutrient give effect as per its presence as low and high. Nitrogen Deficiency Symptoms is young, spindles, stunted plants which reduced growth and yellowing of plants starting with older leaves [12]. So we can say that nutrients give different effect and defect on crops by its low or high value.

F. Decision Tree Approach

There are various algorithms like ID3, C4.5, C5.0 and J48 used in decision tree. We add some concept of it to more generalize the decision tree. In 1993 ID3 (Iterative Dichotomiser 3) was developed by Ross Quinlan. Increased or grown to trees maximum size and a pruning step are usually applied for improve the ability of the tree for unseen data generalisation. Algorithm creates a multiway tree for finding each node in a greedy manner or categorical feature that will yield or create the large information gain for their categorical targets. C4.5 converts the trained trees or the output of the ID3 algorithm into sets of if-then rules. ID3 successor of C4.5, removed the restriction and that features categorised by defining dynamically a discrete attribute based on numerical variables which partitions the continuous attribute values into a discrete intervals set. Rule accuracy is evaluated the order determined in which they would be applied. Step pruning is done and processed by removing a rule's precondition without the accuracy of the rule improves by it. Uses less memory and builds smaller rule sets than C4.5 while being more accurate [17]. Quinlan's latest version C5.0 is under proprietary license.

G. J48 decision tree

J48 implements Quinlan‟s C4.5 algorithm [4] for generating C4.5 pruned or unpruned decision tree. An extension of Quinlan's earlier ID3 algorithm is C4.5. The decision trees J48 can be used for classification. Using the information entropy, J48 builds decision trees from a labelled training data. It uses the fact that each data attribute can be used to make a decision by the data splitting into smaller subsets. J48 which is examines the normalized information gain or difference in entropy that results for splitting the data from choosing an attribute. Highest normalized information gain of attributes is used to make the decision and then the algorithm recurs on the smaller subsets of element. If all instances in a subset belong to the same class then after the splitting procedure stops. Leaf node in the decision tree telling to choose that class which are used. Some times happens that none of the features will give any information gain. So that J48 creates a decision node which higher up in the tree using the expected class value. J48 can handle both discrete and continuous attributes, missing attribute values and attributes of training data which in differing costs. After creation of J48, it provides a further option for pruning trees. For further good information in details, we refer the original publications which in [5].

III. DECISION TREE PROCESS

A. Dataset Collection

Dataset required for this research was collected from soil testing laboratory of micronutrient department; AAU (India) .Large soil samples are taken. Table 1 show attribute selection.

Attribute	**Description**
EC	Electrical conductivity, decisiemen per meter
PH	pH value of soil
OC	Organic Carbon
Fe	Iron
Mn	Manganese
Zn	Zinc
Cu	Copper
B	Boron
K	Potassium
Label	Yellowing Crop, Wilting of Crop, Burning of leaf edges, Inward and outward leaves, Sucking and Other pest

Table -4 Attribute Description

B. Data Formatting

All the data are formatted into an Excel format based on various soil types and relevant related field. All data is in single excel spread sheet.

IV. RESULT

A. WEKA

The WEKA (Waikato Environment for Knowledge Analysis) workbench is the state of- the-art machine learning algorithms and data pre-processing tools. It is an open source collection. Data mining software WEKA is used to determine and predict result if any advantage would be gained in both interpretation of data set and time saving. The application of the data required for WEKA that some of pre-processing to be undertaken. Excel data set converted into .CSV file format to allow them to be applied to WEKA. The data mining platform allowed and uses the number of data interpretations techniques including clustering, classifying and associate routines conducted after the pre processing stage.

B. J48 Result

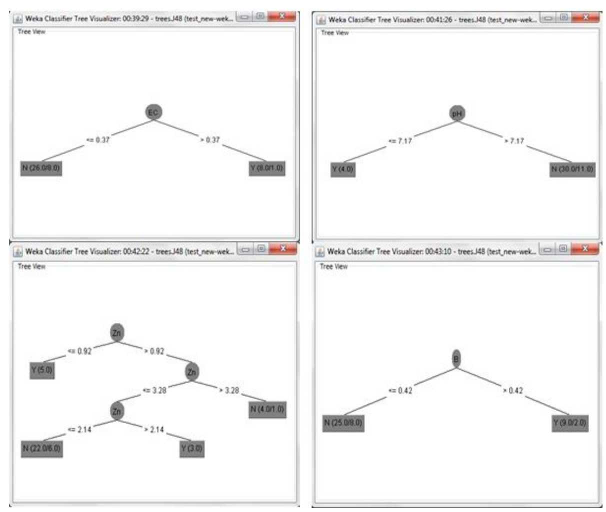
The J48 is slightly modified C4.5 in WEKA. J48 is give higher accuracy result of 91.90 percent for predicts the soil fertility model as per [13]. The soil data set did not require any filtering because of the soil testing data is so accurate [18].So that the limited amount of missing values and the outcomes required by the researchers. The all soil data set is applied to the J48 for classify the soils and established for constructed the model by using a training model for classify the training data set and see the outcomes of the correctly classified instances. Apply data set to J48, it gives better result by experiment with the example of Yellowing Crop which are explain here.

1) WEKA J48 Result: Experiment result on nutrient of Bt Cotton give different ranging values that show yellowing crop disease is available or not. J48 result give 83.74 percent correctly classified instances classified data.

Nutrient	**Yellowing Crop**
EC	N <= 0.37 <Y
PH	Y <= 7.17 < N
OC	Not Related
Fe	Not Related
Mn	Not Related
Zn	Y <= 0.92 < N<= 2.14 < Y <= 3.21 <N
Cu	Not Related
B	N <= 0.42 < Y
K	Not Related

Table -5 Nutrient affect Yellowing Crop deficiency with its availability

2) WEKA J48 Tree of Yellowing Crop:



3) Rating Of Nutrient: Nutrient Ratings of soils for different parameters is followed in soil testing laboratories of Gujarat. This four nutrient rate is taken for more understand that how its rate affect to crop by low or higher value is available in soil.

Element/Parameter	Low	Medium	High
pH	< 6.5 (Acidic)	6.5 - 8.2 (Neutral)	> 8.2 (Alkaline)
EC	< 1.0	1.0 - 3.0	> 3.0
Zn	< 0.5	0.5 – 1.0	> 1.0
B	< 0.2	0.2 – 0.4	> 0.4

Table-6: Nutrient rating of pH, EC, Zn and B

4) Nutrient Relation and Value Dependent

Yellowing crop deficiency is based on shortage of Fe but it 's availability also related to pH and Zn rating. Here ph and Zn is higher rating come up through J48 result. If pH and Zn is higher then it decrease the Fe in soil then after they create this yellowing symptoms on crop. EC is lower as per rating table. B is no harm at 3.0 mg/kg, here rating say that if 0.42 and higher then it may be cause of yellowing crop but less is not affect, but higher rate is 0.40 mg/kg as per rating table so it will be create deficiency related to B as per table 2.

As per this experiment we can find other deficiency like wilting of crop, burning of leaf edges, inward and outward leaves, sucking and other pest are available or not and depends their effect on crop which related with each other.

V. Conclusions

Nowadays virtually harvested along with the crops have to be analyzed and should be used to their full extent with the help of large amount of data. Various decision tree algorithms can be used for prediction of soil fertility. My studies showed that J48 gives 83.74 % accuracy; hence it can be used as a base learner. If we have data about soil, crop eye observation and nutrient functions which give the general characteristic of nutrients effect than we make better prediction model that help to improve production of crop.

Acknowledgment

The author gratefully acknowledges the support of Dr. V. P. Ramani, Micronutrient Project (ICAR), Anand Agriculture University who help me more understanding of soil nutrients.

References

[1] J.W. Doran, Soil health and global sustainability: translating science into practice, Proceedings of International Workshop on Soil Quality as an Indicator of Sustainable Land Management. Goulandris Natural History Museum. Gaia Environmental Research and Education Center, Athens, Greece, 1999.

[2] X.Wang, Z.Gong, "Assessment and analysis of soil quality changes after eleven years of reclamation in subtropical China". Geoderma 81, 1998,pp.339–355.

[3] J.M. Bremner, C.S. Mulvaney, "Nitrogen total. In: Buxton,D.R. (Ed.), Methods of Soil Analysis", Part 2: Chemical Analysis. American Society of Agronomy Inc. And Soil Science Society of America Inc., Madison, WI, 1982, pp. 595–624.

[4] Ross J. Quinlan: Learning with Continuous Classes. In: 5th Australian Joint Conference on Artificial Intelligence, Singapore, 343-348, 1992.

[5] Ross Quinlan (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, CA.

[6] Orchid House, *Micronutrient Requirements of Crops*. [Online]. Available: http://retirees.uwaterloo.ca/~jerry/orchids/nutri.html

[7] G.W. Thomas. Problems encountered in soil testing methods, soil testing and plant analysis. Part 1.Madison, WI: Soil Science Society of America, Spec Pub No 2. 1967:37

[8] G.W. Thomas, W.L. Hargrov. The chemistry of soil acidity. In: F. Adams, ed. *Soil Acidity and Liming*. 2nd ed. ASA Monogr 12. Madison, WI: ASA, 1984.

[9] S.M. Alam. Effects of solution pH on the growth and chemical composition of rice plant. J Plant Nutr 4(3):247, 1981.

[10] A.K.M.S. Islam, D.G. Edwards, C.J. Asher. pH optima for crop growth. Results of a flowing solution culture experiment with six species. Plant Soil 54:339, 1980.

[11] Ratings of soils for different parameters, *Micronutrient*, Soil Testing Laboratory, AAU, Gujarat.

[12] M. Thakura, S. Olafssonb, J-S. Leeb, C.R. Hurburgha, Data Mining for Recognizing Patterns in Food borne Disease Outbreaks, Journal of Food Engineering 97(2), 213–227, 2010.

[13] Jay Gholap, *Performance Tunningof J48 Algorithm for Prediction of Soil Fertility*,

[14] El Dorado Chemical Company, *Roles of the 16 Essential Nutrients in Crop Development*. [Online]. Available: http://www.eldoradochemical.com/fertiliz1.htm.

[15] Dr. Ross H. McKenzie, *Agriculture Rural Development,* Soil Fertility Specialist, Soil Branch, Lethbridge. [Online]. Available: http://www1.agric.gov.ab.ca/$department/deptdocs.nsf/all/agdex713.

[16] Tamilnadu Agriculture University, *Mineral Nutrition*, TNAU-2013. [Online]. Available: http://agritech.tnau.ac.in/agriculture/agri_min_nutri_def_symptoms.html.

[17] Samuel A. Moore, Daniel M. D'Addario, James Kurinskas and Gary M. Weiss, Are Decision Trees Always Greener on the Open (Source) Side of the Fence?, Department of Computer and Information Science, Fordham University, Bronx, NY, USA.

[18] Vamanan R. & Ramar K. (2011) "CLASSIFICATION OF AGRICULTURAL LAND SOILS A DATA MINING APPROACH"; International Journal on Computer Science and Engineering (IJCSE);ISSN: 0975-3397 Vol. 3.

[19] Nutrient Deficiency, NC State University, [Online]. Available: http://www.ces.ncsu.edu/depts/hort/consumer/quickref/fertilizer/nutri_def.html

[20] El Dorado Chemical Company, Roles of the 16 Essential Nutrients in Crop Development. [Online]. Available: http://www.eldoradochemical.com/fertiliz1.htm.

[21] Orchid House, Micronutrient Requirements of Crops. [Online]. Available: http://retirees.uwaterloo.ca/~jerry/orchids/nutri.html

[22] Tamilnadu Agriculture University, Mineral Nutrition, TNAU-2013. [Online]. Available: http://agritech.tnau.ac.in/agriculture/agri_min_nutri_def_symptoms.html.