Disease Prediction on Soil Micronutrients Analysis of Bt Cotton By J48 Classification

[1]Youvrajsinh Chauhan, [2]Jignesh Vania
[1]Computer Engineering, [2] Dept of IT, LJIET
[1]LJ Group of Institutes, Ahmedabad, India

Abstract - **The decision tree is one of the common modelling methods to classify. This is an emerging research field that is experiencing a constant development. In this paper, we present the basic knowledge of 16 soil nutrients of crop and how related with other factors like Soil PH, Cation Exchange, Soil Fertility. We add general concept of decision tree and recent applications that already tested and predict the use in field of agriculture. We make the conceptual model that helps to feature research work in agriculture with use of soil nutrients.**

Index Terms - **Classification, Data Mining, Soil Nutrients, Soil Science, J48**

I. INTRODUCTION

Soil fertility, the health of soil for growing plants, is recognized as a primary constraint to agricultural production in developing countries. 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. 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. 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.

Here we cover 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 soil fertility levels other then NBTree and SimpleCART algorithms [1].

This report is organized as follows. In Section 2, we will try to present the concept and literature review. Section 3, take an experimental data and try to understand process here. Section 4, proposed work and last section 5 we introduce tools and finish with conclusion.

II. STUDDING SOIL NUTRIENTS IN AGRICULTURE

In Agriculture sixteen plant food nutrients are essential for proper crop development 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.

Primary (macro) nutrients are nitrogen (N), phosphorus (P), and potassium (K). They are the most frequently required in a crop fertilization program. All are need in the greatest total quantity by plants as fertilizer. The secondary nutrients are calcium (Ca), magnesium (Mg), and sulphur (S). For most crops, these three are needed in lesser amounts that the primary nutrients. They are growing in importance due to more stringent clean air standards and efforts to improve the environment in crop fertilization programs. The micronutrients are boron (B), chlorine (Cl), cooper (Cu), iron (Fe), manganese (Mg), molybdenum (Mo), cobalt (Co), silicon (Si) and zinc (Zn). These plant food elements are used in very small amounts, but they are just as important to plant development and profitable crop production as the major nutrients [2] [3]. Especially, they work "behind the scene" as activators of many plant functions. All nutrients are needed but it's some limited rating. According to their limited rate, every nutrient has its own different impact on crops. Macro or primary elements are importance for the growth of the plant but micro elements are not less important than those of major elements. Due to deficiency of these elements leaves, branches and fruits may not properly grow and they may even affect the fruits quality as well as production. Micro elements also help in development of enzymes, hormones, chlorophyll and in the absorption of major elements. Nutrient functions, deficiency symptoms, availability are used for agricultural land which is very effectively give impact on soil.

Farmers Problem : Bt Cotton

Deficiency of a nutrient does not directly produce symptoms. Visual evaluation of nutrient stress should be used only as a supplement to other diagnostic techniques (i.e., soil and plant analysis). Nutrient deficiency symptoms may be classified as follows by [4]:

- Crop failure at the seedling stage.
- Severe stunting of plants.
- Specific leaf symptoms appear at varying seasonal times.
- Internal abnormalities like as clogged conductive tissues.
- Delayed or abnormal maturity.
- Differences on yield, with or without leaf symptoms.
- Poor quality of crops or plants with differences in oil, protein, or starch content, and storage quality.

Soil Type

The soils of Gujarat can be broadly classified into nine groups: black soil, mixed red and black soils, residual sandy soils, alluvial soils, saline/alkali soils, lateritic soils, hilly soils, desert soils and forest soils [10].

III. EXPERIMENTAL ANALYSIS

Soil Testing Data Attributes

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
Soil Type	Goradu Soil, Black Soil, Mixed Red and Black Soil, Light Black Soil, Medium Black Soil, Bhatta Soil
Observation	Yellowing Crop, Wilting of Crop, Burning of leaf edges, Inward and outward leaves, Sucking and Other pest

Rating Of Nutrients

Ratings of soils for different parameters as followed in soil testing laboratories of Gujarat [11].

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
OC	< 0.50	0.50 - 0.75	> 0.75
N	< 240	240 - 480	> 480
P	< 28	28 - 56	> 56
K	< 140	140 - 280	> 280
S	< 10	10 - 20	> 20
Fe	< 5	5 – 10	> 10
Mn	< 5	5 – 10	> 10
Zn	< 0.5	0.5 – 1.0	> 1.0
Cu	< 0.2	0.2 – 0.4	> 0.4
B	< 0.2	0.2 – 0.4	> 0.4

DTA Comparison

Classifier	NBTree	SimpleCart	J48
Correctly Classified Instances	1668	1768	1804
incorrectly Classified Instances	332	272	196
Accuracy (%)	85.51	88.44	90.13

IV. IMPLEMENTATION STRATEGY

HTML and CSS

Hypertext Markup Language (HTML) is a markup language for creating web pages or other information to display in a web browser. HTML allows images and objects to be included and that can be used to create interactive forms. From this, structured documents are created by using structural semantics for text such as headings, links, lists, paragraphs, quotes etc.

CSS (Cascading Style Sheets) is designed to enable the separation between document content (in HTML or similar markup languages) and document presentation. This technique is used to improve content accessibility also to provide more flexibility and control in the specification of content and presentation characteristics. This enables multiple pages to share formatting and reduce redundancies.

PHP and Macromedia Dreamweaver

PHP (recursive acronym for PHP: Hypertext Preprocessor) is a widely-used open source general-purpose server side scripting language that is especially suited for web development and can be embedded into HTML.

Macromedia Dreamweaver is a web design and development application that provides a visual editor (colloquially referred to as the Design view) and a code editor with standard features such as syntax highlighting, code completion, and code collapsing as well as more sophisticated features such as real-time syntax checking and code introspection for generating code hints to assist the user in writing code. The Design view facilitates rapid layout design and code generation as it allows users to quickly create and manipulate the layout of HTML elements [12]. Dreamweaver features an integrated browser for previewing developed web pages in the program's own preview pane in addition to allowing content to be open in locally installed web browsers.

MySQL

MySQL is the most popular open source RDBMS which is supported, distributed and developed by Oracle. In the implementation of our web application, we have used it to store user information and soils' data.

WEKA Classifier and Java

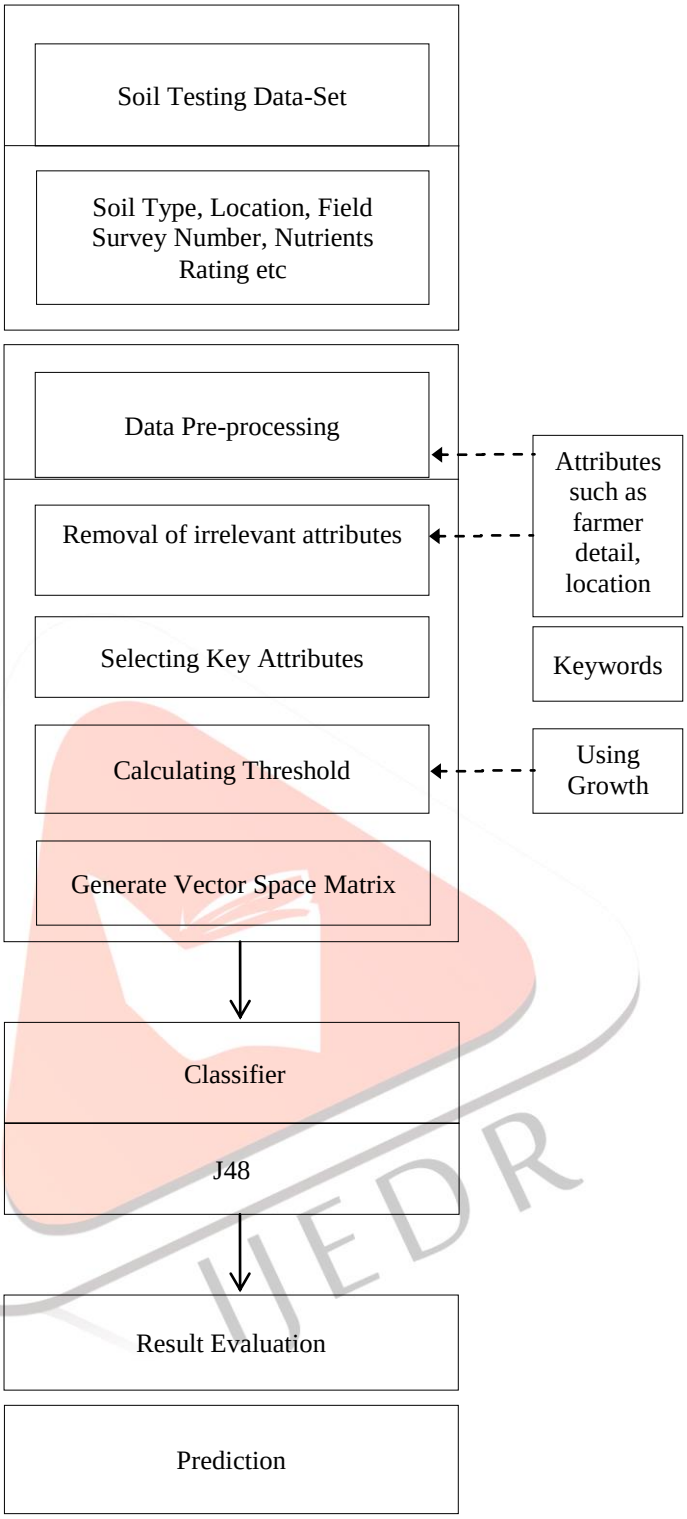
The WEKA (Waikato Environment for Knowledge Analysis) workbench is the state of- the-art machine learning algorithms and data pre-processing tools. It is an open source collection. Data mining software WEKA is used to determine and predict result if any advantage would be gained in both interpretation 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.

JAVA is open source language. Here we use java library with WEKA classifier for getting decision tree output come out by J48 classifier algorithm and then after we can use class method for implementation.

Implementation

We had divided the entire implementation into five stages. In the first stage, information about soil testing data which will apply for prediction diseases. In the second stage, extraneous information was removed from the collected data and the relevant information was fed into a database. The third stage involved applying the J48 algorithms on the training data to obtain decision trees of the algorithms. In the next stage, the test data is applied to the decision trees. The final stage consisted of developing the front end in the form of a web application.

The stages of implementation are depicted here.



Soil Testing Database

We were provided with training dataset consisting of information about soil testing data. The data was in the form of Microsoft Excel 2007 in CSV format and had details of each soil such soil type, location and field survey no, nutrients rating etc. For ease of performing data mining operations, the data was filled into MySQL database.

Data Preprocessing

Once we had details of all soil testing data, we then segmented the training dataset further, considering various feasible splitting attributes, i.e. the attributes which would have a higher impact on the prediction. For instance, we had considered 'location' as splitting attributes.

A snapshot of soil testing database is shown in Figure 2. Here, irrelevant attributes such as soil location, type, and farmer detail etc. had been removed. Here we pre-processed data using rating of nutrients. For example PH is medium in between value of 6.5 to 8.2 and respectively low and high of less than 6.5 and greater than 8.2. I use separate code for preprocess shown in Figure 1 and 2.

Figure 1.Preprocessed Model

```
$EC = $r['EC'];
if ($EC<1.0) { $EC='LOW';} elseif ($EC>1.0 && $EC <3.0) { $EC='MEDIUM';} else { $EC='HIGH';}
$PH = $r['PH'];
if ($PH<6.5) { $PH='LOW';} elseif ($PH>6.5 && $PH <8.2) { $PH='MEDIUM';} else { $PH='HIGH';}
$OC = $r['OC'];
if ($OC<0.50) { $OC='LOW';} elseif ($OC>0.50 && $OC <0.75) { $OC='MEDIUM';} else { $OC='HIGH';}
$Fe = $r['Fe'];
if ($Fe<5) { $Fe='LOW';} elseif ($Fe>5 && $Fe <10) { $Fe='MEDIUM';} else { $Fe='HIGH';}
$Mn = $r['Mn'];
if ($Mn<5) { $Mn='LOW';} elseif ($Mn>5 && $Mn <10) { $Mn='MEDIUM';} else { $Mn='HIGH';}
$Zn = $r['Zn'];
if ($Zn<0.5) { $Zn='LOW';} elseif ($Zn>0.5 && $Zn <0.1) { $Zn='MEDIUM';} else { $Zn='HIGH';}
$Cu = $r['Cu'];
if ($Cu<0.2) { $Cu='LOW';} elseif ($Cu>0.2 && $Cu <0.4) { $Cu='MEDIUM';} else { $Cu='HIGH';}
$B = $r['B'];
if ($B<0.2) { $B='LOW';} elseif ($B>0.2 && $B <0.4) { $B='MEDIUM';} else { $B='HIGH';}
$K = $r['K'];
if ($K<140) { $K='LOW';} elseif ($K>140 && $K <280) { $K='MEDIUM';} else { $K='HIGH';}
```

Figure 2. Pre-processed Soil Database

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Farmer	SOIL TYPE	EC	pH	OC	Fe	Zn	Cu	B	K	Yellowing	Wilting	of Burning o	Inward cr	Outward (	Sucking pe	Other pes	Any disea
2	Large	Goradu so	Low	Medium	Low	High	Medium	High	Medium	High	Y	Y	N	Y	Y	Y	Y	N
3	Large	Goradu so	Low	Medium	Low	High	High	High	Medium	High	Y	N	N	N	N	Y	N	Y
4	Medium	Goradu so	Low	Medium	Low	High	High	High	Low	High	N	N	N	N	N	Y	N	N
5	Medium	Goradu so	Low	Medium	Medium	High	High	High	Medium	High	N	Y	N	Y	N	Y	N	N
6	Large	Black Soil	Low	Medium	Medium	High	High	High	Medium	High	N	Y	N	Y	N	N	N	N
7	Large	Black Soil	Low	Medium	Medium	High	High	High	High	High	N	Y	N	Y	N	N	N	N
8	Medium	Black Soil	Low	Medium	Medium	High	High	High	High	Medium	N	Y	N	Y	N	N	N	N
9	Medium	Black Soil	Low	Medium	Medium	High	High	High	High	High	N	Y	N	Y	N	Y	Y	Y
10	Large	Goradu so	Low	Medium	Low	High	High	High	Medium	High	N	Y	N	N	N	N	N	N
11	Large	Goradu so	Low	Medium	Medium	High	High	High	Medium	Low	N	Y	N	N	N	N	N	N
12	Medium	Goradu so	Low	Medium	Low	High	High	High	Low	Low	Y	N	N	N	N	N	N	N
13	Medium	Goradu so	Low	Medium	Low	High	High	High	Low	High	Y	Y	N	N	N	N	N	N
14	Large	Black Soil	Low	Medium	Medium	High	High	High	Low	High	Y	N	N	N	N	N	N	N
15	Large	Black Soil	Low	Medium	Low	High	High	High	Low	High	Y	Y	N	N	N	N	N	N
16	Medium	Black Soil	Low	Medium	Low	High	High	High	Low	High	Y	N	N	N	N	N	N	N
17	Medium	Black Soil	Low	Medium	Medium	High	High	High	Low	High	N	Y	N	N	N	N	N	N
18	Large	Goradu so	Low	Medium	Medium	High	Medium	High	Low	High	N	Y	N	N	N	N	N	N
19	Large	Goradu so	Low	Medium	Low	Medium	Medium	High	Medium	Medium	N	Y	N	N	N	N	N	N
20	Medium	Black Soil	Low	Medium	Medium	High	Medium	High	High	High	Y	Y	N	N	Y	Y	N	N
21	Medium	Black Soil	Low	Medium	Medium	High	High	High	High	High	Y	Y	N	Y	Y	Y	N	N

Data Processing Using WEKA Classifier and PHP

This is tree view of yellowing crop which we used if-else ladder for prediction.

Figure 3. Tree View

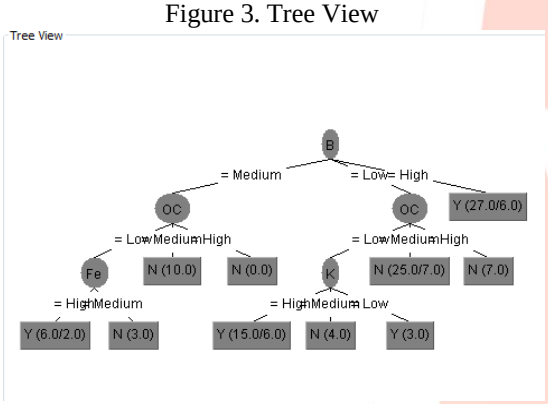
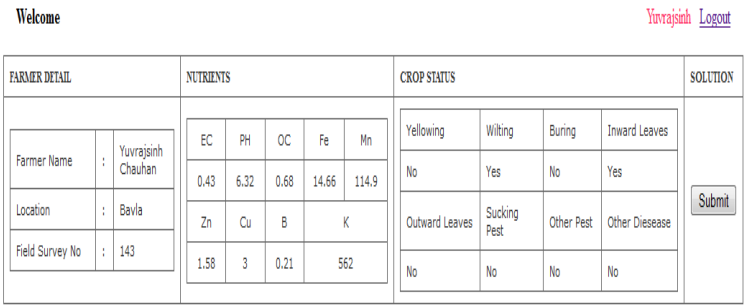


Figure 4 Prediction



V. CONCLUSION

There are a growing number of applications of decision tree in agriculture and growing amount of data which are currently available from many resources. This is relatively a modern research field and it is expected to grow in future. There is lot of work to be done on this emerging area. Here we get a basic knowledge of agriculture nutrient and how it will be classified and leads toward decision making on Bt Cotton crop farming and helps to find current status of crop on disease with J48 classification approach that provide an effective division for farmland fertility level partition and use of nutrient availability that may give the better growth, production, remove disease and may help for find true decision for Bt Cotton crop farming in agriculture.

REFERENCES

[1] Jay Ghopal, Performance Tunning Of J48 Algorithm For Prediction Of Soil Fertility, Asian Journals of Computer Science and Technology, Vol-2, 2012.

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

[3] 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.

[4] Jiejun HUANG, Yanbin YUAN, Wei CUI, Yunjun ZHAN, An Approach Based on Decision Tree to Agricultural Land Grading, International Conference On Computer Design And Applications, 978-1-4244-7164-51$26.00 © 2010 IEEE.

[5] He, Cao Liying, Zhang Xiaoxian, Li Dexin, A Study On Evaluation of farmland fertility levels based on optimization of the decision tree algorithm, 2nd International Conference on Computer Science and Network Technology, 978-1-4673-2964-4/12/$31.00 ©2012 IEEE.

[6] CAO Weixing, TANG Liang, ZHU Yan, PAN Jie, LI Weiguo, CHEN Binling, Development of Growth Model-based Decision Support System for Crop Management, 0-7695-2851-1/07 $20.00 © 2007 IEEE, DOI 10.1109/PMA.2006.51.

[7] Jun Wu, Anastasiya olesnikova, Chi-Hwa Song, Won Don Lee, The Development and Application of Decision Tree for Agriculture Data, 978-0-7695-3579-1/09 $25.00 © 2009 IEEE, DOI 10.1109/IITSI.2009.10.

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

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

[10] Soil of Gujarat, Booklet No 360, Soil Science- SSS -5.

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

[12] Adobe Dreamweaver [Online]. Available: http://en.wikipedia.org/wiki/Adobe_Dreamweaver.