Data Mining in Soil & Plant Nutrient Management, Recent Advances and Future Challenges in Organic Crops

S. Jeyalaksshmi, V. Rama, G. Suseendran

Abstract--- Data mining procedures are colossally regarded in the examination field of cultivating. The cultivating factors atmosphere, deluge, soil, pesticides, and fertilizers are the basic careful perspective to raise the age of yields. The key piece of agribusiness is Soil for yield creating. Examination of soil is a basic bit of soil asset the administrators in development. The earth examination is especially significant for cultivators to observe which sort of harvests to be made in a specific soil condition. The rule focal point of this work is to look at soil supplements utilizing data mining gathering techniques. A broad enlightening record of soil supplements status database was accumulated from the Department of Agriculture, Cooperation and Farmers Welfare. The database contains the estimation of soil supplements for each and every different state. This work takes some district of Tamil Nadu in India to look at the soil supplements. Specific sort's earth has a different variety of improvements. This paper picks Nitrogen, Phosphorus, Potassium, Calcium, Magnesium, Sulfur, Iron, Zinc, and so on, supplements for investigating the earth upgrades utilizing Hybrid procedure of Neural framework. The execution of the portrayal computations is taken a gander at subject to the going with two factors: accuracy and execution time.

Keywords--- Data Mining, Agriculture, Soil Nutrients, Classification, Hybrid Neural Network.

I. INTRODUCTION

Horticulture is for the most part reliant on the soil quality yet over the long haul increasingly agrarian generation brings about the loss of supplements present in the soil. It requires recognizing methods that will back off this disposal of supplements and furthermore will restore the required supplements with the soil, so continue getting high caliber and the great amount of crop productions [1].

In horticulture, great soil wellbeing implies the capacity of soil to forces physical, compound and organic exercises for reliable profitability of high product yield. Great nature of soil guarantees us for maintenance and arrival of water and supplements, upgrade and steady root development while keeping up biotic condition, giving the normal outcome and opposes foulness [2]. Soil is basic for vegetation. It is made out of solids (minerals and natural issues), fluids (water and broke down substances), and gases (for the most part oxygen and carbon dioxide) and contains living creatures. Every one of these parts offer their physical and synthetic properties.

Dealing with the dirt legitimately is fundamental so as to safeguard its ripeness, get better yield and regard nature. Testing the dirt on the contrary hand might be an ought to in order to oversee it legitimately.

A dirt test is the examination of a dirt guide to find supplement substance, sythesis, and distinctive characteristics. Soil tests are ordinarily performed to choose wealth and show deficiencies that ought to be restored [3]. Soil supplements examination is important for the agriculturist to make sense of which sort of respects be created in a particular soil condition.

Concentrate on order of soil extravagance rate utilizing J48, Naïve Bayes, and Random timberland calculation. J48 calculation gives an ideal result over different calculations. Choice tree structure by J48 calculation helps the cultivator and leaders to distinguish the dirt lavishness rate and on the reason of supplements found in the dirt example, diverse manures can be recommended.

The dirt has 3 assortments of qualities: physical, substance, and natural. This paper for the most part centers around concoction qualities. Synthetic properties of soil include the administration of soil supplements at the first essential dimension. Mineral supplements inside the dirt have either a positive or a negative charge. The supplements with an electric charge (counting Ca, magnesium, potassium, ammonium, and sodium) are held by contrarily charged mud particles and natural issue. The dirt supplies the accompanying basic supplements for legitimate plant development and creation. These supplements will begin from weatherworn minerals or from decaying natural issue. The supplements are characterized into three structures Primary, Secondary and Micro. The Nutrients and their plant accessible structures are

- **Primary Nutrients:** Nitrate, Ammonium Phosphorus and Potassium
- **Secondary Nutrients:** Calcium, Magnesium
- **Sulfur Micronutrients:** Iron, Manganese, Boron, Molybdenum, Copper, Zinc, Chlorine, Nickel and Cobalt.

The primary nutrients square measure employed in the best quantity by the plant. Analyzing nutrients is one of the main topics in the agriculture field. It is most helpful for agriculturists to make sense of which kind of respects be created in a particular soil condition. This paper focuses on the classification of the soil nutrients analysis based on the selected nutrients for some district (Ariyalur, Coimbatore, Karur, Salem, Thanjavur, and Trichy) in Tamil Nadu.

Manuscript received September 16, 2019.

S. Jeyalaksshmi, Assistant Professor, School of Computing Sciences, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai. T.N, India. (e-mail: jlakshmi.scs@velsuniv.ac.in)

V. Rama, Research Scholar, School of Computing Sciences, VISTAS, Chennai. T.N, India. (e-mail: rama.ve@gmail.com)

Dr.G. Suseendran, Assistant Professor, School of Computing Sciences, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai. T.N, India. (e-mail: suseendar_1234@yahoo.co.in)

Retrieval Number: B10350982S1119/2019©BEIESP
DOI: 10.35940/ijrte.B1035.0982S1119

213

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

II. OBJECTIVES

The transitional period for new natural tasks can be the most requesting regarding soil ripeness the executives. This is because of the benefits of soil building and soil natural issue improvement haven't by the by been finished in an exceedingly change field, yet the cultivator is restricted to just a couple of dissolvable manure materials. Cultivators gain understanding amid the transitional period, and as the dirt natural issue manufactures, its advantages are reflected in improved soil richness. Affirmed materials records have been created by the NOP and incorporate a wide scope of materials to enhance crop dietary needs. Materials from natural or common sources may contain squander vegetables, Crops, seeds, Green fertilizer and Vermi composting that may collect to harmful dimensions on a given field.
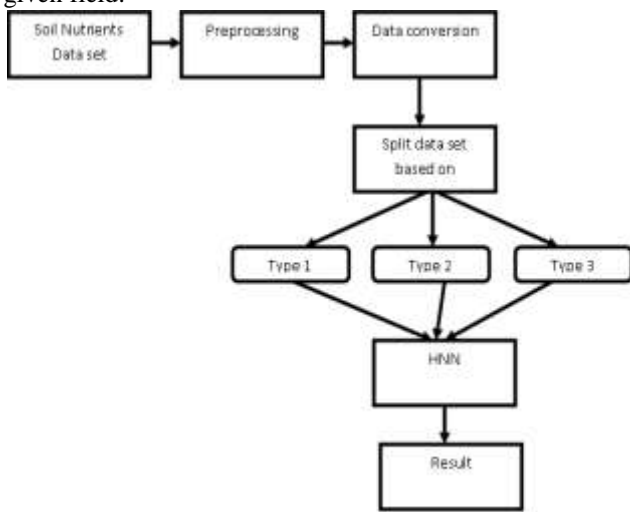

Fig. 1: Proposed Work Flow

III. RELATED WORK

Baskar et al.,[4] examination the soil information exploitation distinctive algorithms and statement technique. A report is exhibited exploitation completely.

Jay Gholap [5] predicts soil extravagance abuse call tree equation. In [6], the creator determined soil qualities and inspected soil data abuse characterization strategies. Soil properties, for instance, pH esteem, Electrical Conductivity (EC), Potassium, Iron, Copper, and so forth ordered utilizing arrangement calculations [7] use KNN, Naïve Bayes and J48 for breaking down soil information. This examination recommends the manure dependent on the dimension of supplements found in the dirt test set.

Bhuyar [2] concentrate on order of soil extravagance rate utilizing J48, Naïve Bayes, and Random timberland calculation. J48 calculation gives an ideal result over different calculations. Choice tree structure by J48 calculation helps the cultivator and leaders to distinguish the dirt lavishness rate and on the reason of supplements found in the dirt example, diverse manures can be recommended.

Hemageetha et al.,[8] investigate whether the Salem district soil is appropriate or nonsuitable for crops development in view of pH value using data mining classification techniques. The outcome demonstrates that the accuracy of J48 is high compared to Naive Bayes, Jrip, and BayesNet. And it also shows that the major part of the

Salem locale soil is appropriate for the development of many harvests.

Bhargavi et al.,[9] contended GATree, Fuzzy Classification rules and Fuzzy C-Means calculation for ordering soil surface in cultivation soil information. Characterization dependent on Fuzzy principles gives much execution than GATree [10], calculations are analyzed. Characterization calculation improved it is proficiently grouped into most extreme number of cases contrasting and the other two.

Ramesh et al., [11] clarifies correlation of various classifiers and the result of this examination could improve the administration and frameworks of soil clients, all through a huge fields that incorporate farming, cultivation, natural and land use the executives. Dildarkhan et al.,[12] gives an examination of the dirt information utilizing diverse arrangement calculations and estimating techniques. Soil testing research focuses examination the dirt and gives the example informational collection. It will require an extensive proportion of time to describe the dirt datasets physically. Shivnath et al.,[13] an examination of soil properties utilizing Back Propagation Network. This examination utilizes the angle plunge calculation in preparing the examples. Back Propagation calculation creates the better outcome.

IV. SOIL NUTRIENTS ANALYSIS

Data Mining is fundamental to choose the agrarian related substances, for instance, soil readiness, yield desire and soil supplements examination. This fragment examination soil supplements subject to blend request computation. Fig 1 exhibits the designing of proposed work.

Overview of Dataset

To begin with any information mining issue, at first unite every one of the information. The informational collection contains 13 properties.

Methodology

In progression is gathered information was pre-handled. A few records have missing property estimations, that records were expelled from the dataset. In the information transformation step, the pre-prepared information was changed over dependent on the supplements esteems.

Table 1: Soil Nutrients Data set Description

Attribute	Description
District	The data was collected from six district of Tamil Nadu in India.
pH	pH value of Soil Data
EC	Electrical Conductivity
OC	Organic Carbon
N	Nitrogen
P	Phosphorus
K	Potassium
S	Sulfur
Zn	Zinc
Fe	Iron
Cu	Copper
Mn	Manganese
B	Boron

Table 2: Nutrients Levels

Level\nAttribute	Low | Deficient	Medium | Moderate	High | Sufficient
pH	< 6.5 (Acidic)	6.5 – 7.5 (Neutral)	> 7.5 (Alkaline)
EC	< 1.0 (Non Saline)	1.0 – 3.0 (Slightly Saline)	> 3.0 (Saline)
OC	< 0.5	0.5 – 0.75	> 0.75
N	< 280	280 – 450	450
P	< 11	11 – 22	> 22
K	< 118	118 – 280	> 280
S	< 10	10 – 15	> 15
Zn	< 1.2	1.2 – 1.8	> 1.8
Fe	< 3.7	3.7 – 8.0	> 8.0
Cu	< 1.2	1.2 – 1.8	> 1.8
Mn	< 2.0	2.0 – 4.0	> 4.0
B	< 0.46	0.46 – 1.0	> 1.0

Table 2 demonstrates the incentive for three dimensions of supplements.

After information change, he large scale and miniaturized scale supplements are part into three sorts. Convolutional Neural Network In deep learning, an undemanding probabilistic classification technique is CNN classification.

It relies upon the Bayes hypothesis with autonomy characteristics [10].

The class names are anticipated dependent on the likelihood. For order, a little measure of preparing information is relied upon to anticipate the class marks. [14].

Hybrid Algorithm

In deep learning, CNN may be a class of neural networks, mostly used to analyze visual imagination. Half overlap nominatively, the receptive fields of assorted neurons cowl the entire visual field.

RNN can be an artificial neural network class where node-to-node connections create a direct graph on a temporal sequence. Not like feed-forward neural networks, RNNs can use input sequences of their internal state (memory).

Algorithms

/Input: Images
/Output: accuracy of recognition

1. Load images and labels of the input.
2. Split each group into the corresponding image number
3. Load the pre-trained HNN.
4. Remove noise & Enhance the given image
5. Divide the image sets into data training and testing.
6. Extract characteristics from deeper layers
7. Get all the training tags from training set
8. Train a different class HNN classifier using the training features.
9. Extract characteristics from test set
10. To find the test label set, use the trained classifier.
11. Get the known test set tags
12. Tabulate the results with a matrix of confusion.
13. Oh, thirteen. Convert the matrix of confusion to percentage points.
14. Show the information

Table 3: Sample Nutrient Data set

District	pH	EC	OC	N	P	K	S	Zn	Fe	Cu	Mn	B
Ariyalur	8.30	0.26	0.54	43.2	17.6	789	10.5	2.45	2.45	3.56	3.67	3.4
Ariyalur	5.10	0.17	0.40	60.60	16.25	60	17.50	0.53	5.23	0.17	3.38	1.0
Coimbatore	6.50	0.38	0.18	231.0	18.0	400	8.10	0.80	5.10	2.20	2.60	0.30
Coimbatore	7.80	0.18	0.18	173.0	13.6	284	17.24	1.0	5.30	1.40	2.60	0.41
Karur	7.90	0.6	0.35	80.00	2.70	189	8.10	0.84	1.40	0.02	1.08	1.10
Karur	8.10	1.50	0.15	45.00	2.70	233	0.11	0.10	2.60	0.03	1.70	1.10
Salem	7.90	0.10	0.28	162	7.50	315	29.2	0.23	10.5	0.72	7.23	2.30
Salem	7.80	0.10	0.11	189	10	378	33.20	0.32	13.5	0.74	0.32	1.0
Thanjavur	6.90	0.15	0.75	205	75	250	25	1.30	6.80	1.80	2.0	1.0
Thanjavur	7.70	0.17	0.69	189	75	300	20.60	1.60	7.40	0.90	2.0	1.0
Trichy	7.0	1.10	0.25	179	53.0	195.0	12.20	1.06	5.72	0.96	3.48	0.10
Trichy	6.80	0.40	0.19	147	25.0	185.0	12.60	0.95	6.27	0.91	5.76	0.10

Table 4: Converted Data set

District	pH	EC	OC	N	P	K	S	Zn	Fe	Cu	Mn	B
Ariyalur	H	L	M	L	M	H	M	H	L	H	M	H
Ariyalur	L	L	L	L	M	L	H	L	M	L	M	M
Coimbatore	L	L	L	L	M	H	L	L	M	H	M	L
Coimbatore	H	L	L	L	M	H	H	L	M	M	M	L
Karur	H	L	L	L	L	M	L	L	L	L	L	H
Karur	H	M	L	L	L	M	L	L	L	L	L	H
Salem	H	L	L	L	L	H	H	L	H	L	H	H
Salem	H	L	L	L	L	H	H	L	H	L	L	M
Thanjavur	M	L	H	L	H	M	H	L	M	H	M	M
Thanjavur	H	L	M	L	H	H	H	M	M	L	M	M
Trichy	M	M	L	L	H	M	M	L	M	L	M	L
Trichy	M	L	L	L	H	M	M	L	M	L	H	L

Table 5: Comparative analysis of classifiers based on pH, EC, OC, N

Algorithm	Correctly Classified Instance	Incorrectly Classified Instance	Accuracy	Execution Time in ms(Milli sec)
CNN	2067	26	93%	1200
RNN	2667	29	91.5%	1600
HNN	2987	1	99.02%	700

V. EXPERIMENTAL RESULT

The collected data set was pre-processed and some records are removed from the data. After preprocessing step, the data set was converted into Low, Medium and High based on nutrients level mentioned in Table 2.The converted data set is shown in Table 4. Apply CNN, RNN and Hybrid classification algorithm to classify the soil nutrients as Very High, High, Medium, Low and Very Low. The comparative analysis of classifiers is given in Table 5. Fig 3 shows the execution time of classification algorithms with three types of nutrients. Fig 2 shows the accuracy rate of classification algorithms with three types of nutrients. It is observed from Fig 3 and Fig 2 that the Hybrid algorithm is able to classify the dataset in less time with better classification accuracy rate.
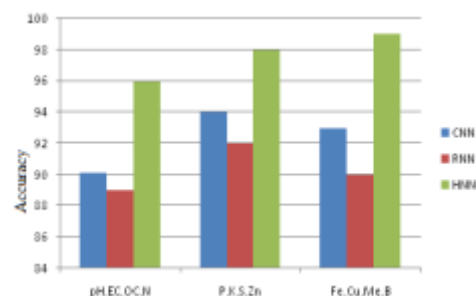

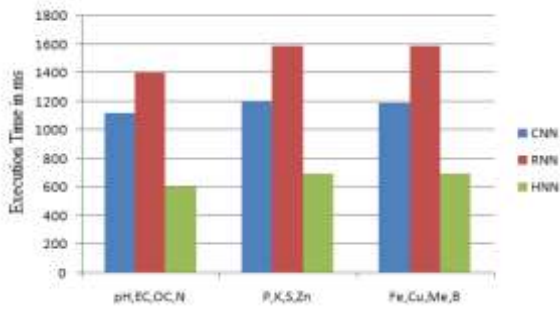**Fig. 2: Accuracy for classifiers**

Fig. 3: Execution Time of classifiers

VI. CONCLUSION

Data mining techniques in cultivation will help the agriculturist to enhance the crop productivity. The yield production is chiefly based on the soil nutrients. The investigation of soil nutrients present whether which type crop can be refined in specifically soil. This paper proposed an investigation of soil nutrients using distinctive algorithms. The comparative examination of three classification algorithms like Naïve Bayes, Decision and hybrid algorithm was anticipated. Hybrid classification algorithm gives enhanced result for this dataset and is correctly classified into maximum number of instances comparing with the other two. Hybrid algorithm can be suggested to dissect the soil nutrients like vegetables, Crops, seeds, Green manure and Vermi composting.

REFERENCES

1. Peter Gruhn, Francesco Goletti, and Montague Yudelman (2000), " Integrated Nutrient Management, Soil Fertility, and Sustainable Agriculture: Current Issues and Future Challenges" Food, Agriculture, and the Environment Discussion Paper 32,2000
2. Bhuyar V. (2014), "Comparative analysis of classification techniques on soil data to predict fertility rate for Aurangabad District", International Journal of Emerging Trends and Technology in Computer Science. 2014 Mar Apr; 3(2):200–3.
3. "Soil test", Wikipedia, February 2012
4. S.S.Baskar L.Arockiam S.Charles (2013), "Applying Data Mining Techniques on Soil Fertility Prediction", International Journal of Computer Applications Technology and Research Volume 2–Issue 6, 660 -662, 2013
5. Jay Gholap (2012), "Performance Tuning of J48 Algorithm for Soil Fertility" ,2012. Asian Journal of Computer Science and Information Technology 2: 8 (2012) 251– 252
6. Gholap J, Lngole A, Gohil J, Shailesh, Attar V.(2012), "Soil data analysis using classification techniques and soil attribute prediction", 2012 Jun; 9(3):1–4
7. Suman, Bharat Bhushan Naib (2013), "Soil Classification and Fertilizer Recommendation using WEKA" ,IJCSMS International Journal of Computer Science & Management Studies, Vol. 13, Issue 05, July 2013
8. Hemageetha, G.M. Nasira (2016), "Analysis of Soil condition Based on pH value Using Classification Techniques", IOSR Journal of Computer Engineering (IOSR-JCE), Volume 18, Issue 6, Ver. III (Nov.-Dec. 2016), PP 50-54
9. Bhargavi P. & Jyothi S. (2011), "Soil Classification Using Data Mining Techniques: A Comparative Study", International Journal of Engineering Trends and Technology
10. Rajeswari V, Arunesh K. (2016), "Analysing soil data using data mining classification techniques", Indian Journal of Science and Technology. 2016 May; 9(19).
11. V.Ramesh and K. Ramar,(2011) "Classification of Agricultural Land Soils: A data Mining approach",Volume:6, Issue: 3,Sep. 2011
12. Dildarkhan T. Pathan, Pushkar D. Joshi, Prof. S. U. Balvir,(2014), "Prediction of soil Quality for Agriculture", IRJSSE International Research Journals of sustainable Science & engineering, Vol. 2, Issues 3, 2014.
13. Shivnath Ghosh, santanu koley,(2014), "Machine Learning for Soil fertility and Plant Nutrients Management using Back Propagation Neural Networks", International Journal on Recent and Innovation Trends in Computing and Communication, Vol. 2, Issues 2, 2014.
14. Narain B. (2011), "Study for Data Mining techniques in classification of agricultural land soils", Journal of Advanced Research in Computer Engineering. 2011 JanJun; 5(1):35–7.
15. Ponmani S, Vidhu Priya P (2017), "Classification Algorithms in Data Mining –A Survey", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 6, Issue 1, January 2017.