ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

An Agri Decision Support System Using Soil and Weather Analysis in Coimbatore District

R.Preetha[1], B.Murugesakumar[2]

M.Phil Scholar, Dept. of Computer Science, Dr.SNS Rajalakshmi College of Arts & Science, Coimbatore, India [1].

Head of the Department, Dept. of Computer Application [UG], Dr.SNS Rajalakshmi College of Arts & Science, Coimbatore, India [2]

ABSTRACT: Agriculture is the backbone of our country. Analysing and managing agricultural activities may help to improve the production and crop. But the current weather and others factors are affecting agriculture activities, from the tool the agriculturist should be able to make in-season decisions according to the location, soil type and current weather.This paper attempts to provide a new decision support system for agricultural activities, which is named as Agro Analyst. The system proposed a novel approach to automating soil classification from agriculture land and provides best decision support based on different factors. This approach exploits the features of a soil using machine learning algorithms. Knowledge of the soil features is then used to learn the composition of the field and its soil horizons. This proposal uses soil classification in various locations in Coimbatore district. There are several existing algorithms proposed for soil classification and agricultural activities. For effective classification and data analysis along with the decision support system, the C5 algorithm is used in agricultural data.

KEYWORDS: Data mining; Classification; C5 algorithm; Decision tree; Soil type; Weather detail

I. INTRODUCTION

Data mining involves the use of erudite data analysis tools to discover previously unidentified, suitable patterns and relationships in large data sets. Data mining tools can include statistical models, machine learning methods such as neural networks or decision trees, and mathematical algorithms. As a result data mining consists of more process. This performs analysis and prediction than collecting and managing data. The objective of data mining is to identify valid, potentially useful, novel and understandable correlations and patterns in existing data. Finding useful patterns in data is known by different names (e.g., knowledge extraction, information discovery, information harvesting, data archeology, and data pattern processing) [1].

Agriculture plays a vital role in India's economy. 54.6% of the population is engaged in agriculture and allied activities (census 2011) and it contributes 17.4% to the country's Gross Value Added (current price 2014-15, 2011-12 series). Given the importance of agriculture sector, Government of India took several steps for its sustainable development. Steps have been taken to improve soil fertility on a sustainable basis through the soil health card scheme, to provide improved access to irrigation and enhanced water efficiency through different schemes and to support for creation of a unified national agriculture market to boost the incomes of farmers. Due to the deficient rainfall as well as unseasonal rains and hailstorms, agricultural production in 2014-15 is estimated to be lower than that in 2013-14, a year of record production. Agriculture is the major work in India. This is also being a backbone of the country, which contributes 10-16% GDP (Gross Domestic Product) to the Indian economy [2].

The current study involved with the Coimbatore district data set, so the study initially brings the basic details of Coimbatore such as Weather, Soil, Land type, Crops and Plant details. The Western Zone of Tamil Nadu comprises of 21 revenue taluks carved out of 7 districts *viz*., Coimbatore, Erode, Namakkal, Karur, Dindigul, Madurai and Theni [2]. It comprises of all the revenue taluks of Coimbatore and Erode districts, Thiruchengodu of Namakkal, Karur and Manapparai of Karur, Nilakottai and Palani of Dindigul, Usilampatti of Madurai and Uthamapalayam and Periyakulam of Theni districts. The proposed system handles the Coimbatore district soil, weather and agriculture details. Following table shows the statistics of total number of talukas& villages coming under Coimbatore districts along with crop area in respective taluka.[3]

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

No.	District/Taluk	No. of taluks	No. of blocks	No. of villages	Area sq.km
1.	Coimbatore	1	5	68	1330
2.	Mettupalayam	1	1	71	697
3.	Avinashi	1	2	53	665
4.	Palladam	1	5	132	1519
5.	Udumalpet	1	3	55	1418
6.	Pollachi	1	4	86	1840
Total		6	20	465	7469

Table 1.0 Geographical details of Coimbatore District

In Coimbatore district, Loamy soil, clayey soil and Calcareous black cotton soil are found.

Type of Soil	Places in Coimbatore District
Moderately sallow Non-calcareous well drained, sandy loam, red soil	Karamadai
	Sulur
	Anamalani
	Pollachi (N & S)
	Annur(North Part)
Deep, calcareous, Moderately drained, clay, very dark grayish brown soil	P.N.palayam
shallow, calcareous, well drained, sandy clay loam, yellowish red soil	Kinathukadavu
	Sulthanpet,
	Pollachi(N& S)
Moderated shallow, calcareous, well drained, sandy clay loam, reddish brown soil	Thondamuthur
	SS. Kulam
	Perur
	Madukarai
	Karamadai(patches)
Deep, calcareous, moderately drained, clay loam, very dark grayish brown soil	Madukkarai
	Perur (Eastern part)
	Pollachi(S) Western part
	Sulur(Central, western part)
Deep, moderately, well drained, clay loam, reddish brown soil	Annur
Deep, Non calcareous, excessively drained, gravelly sand, dark brown soil	P.N.Palayam(North part)
	Karamadai
	Madukkari
very deep, calcareous, well drained, sandy clay loam, dark yellowish brown soil	Thondamuthur (west part)
	P.N.Palayam (west & north)
Deep, calcareous, gypsum rich ,	Pollachi(south, eastern part)

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

poorly drained, clay, very dark grey soil	Kinathukadavu(ventral & eastern part)
	Sulthanpet(central part) Sulur(patches, southern part)

Table 2.0 Soil types in Coimbatore district

The above details are collected from Tamil Nadu Agriculture University. The table represents the available soil type in Coimbatore district. This study takes almost all series and area for soil classification and agricultural activity prediction.
The following table represents the average rain fall of Coimbatore district for every season. The rain fall is measured by milli meter. The rain fall in an area determines the crop pattern. Some plants need more rain and some need average. The following table and figure shows the average rainfall of Coimbatore district on every season [4].

Season	Rain fall (mm)
Winter	10.80
Summer	141.60
South West monsoon	164.60
North East Monsoon	350.00
Grand Total	677.00

Table 3.0 Rain fall details of Coimbatore for every season

Major horticulture crops cultivated in this district are fruits crops like mango and banana, vegetables like tomato, brinjal, bhendi and onion, spices like turmeric and flowers like tube rose, and jasmine.

S.No.	Major field crops cultivated	Area ('000 ha)				
		Kharif		Rabi		Total
		Irrigated	Rainfed	Irrigated	Rainfed	
1	Sorghum	0.8	23.4	0.4	9.8	34.4
2	Groundnut	0.6	5.5	1.0	1.0	8.2
3	Maize	2.0	0.3	2.3	1.0	5.6
4	Cow pea	0.2	3.6	-	0.9	4.8
5	Rice	1.6		0.9		2.5
6	Sugarcane	2.5				2.6
7	Horsegram	-	0.8	-	1.0	1.9
8	Bengalgram	-	0.3	-	1.3	1.6
9	Greengram	-	1129	-	431	1.5
10	Cotton	0.7	501			1.2
11	Blackgram	0.0	0.7	0.0	0.3	1.1

Table 4.0 Horticulture crops cultivated in Coimbatore district

II. RELATED WORK

Following section explain proposed system and agroanalyst using data mining process, the proposed system introduces a set of data mining techniques to find an optimal decision on agriculture to increase the crop yielding.

A. *PROPOSED SYSTEM*

- Analyzes the weather detail
- Predicts the weather for next 5 days
- Find the soil types by its features
- Identifies the best plant which is suitable for the soil and current weather
- Brings the appropriate decisions to the user based on the above feature

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

B. *AGROANALYST USING DATA MINNG PROCESS*

[5]The proposed system AgroAnalyst has divided into four phases. At each phases all the algorithm were used to analyse the weather, soil datasets. In first phase, the precipitation (rainfall) , soil and vegetation patterns were generated and its relationships with other weather parameters were found. In second phase, the output of weather forecasting variable and soil type was classified. In third phase, the testing method adopted for this research was percentage split that train on a percentage of a dataset, cross validate on it and test on the remaining percentage, and the final stage involves the suggestion process using DSS.

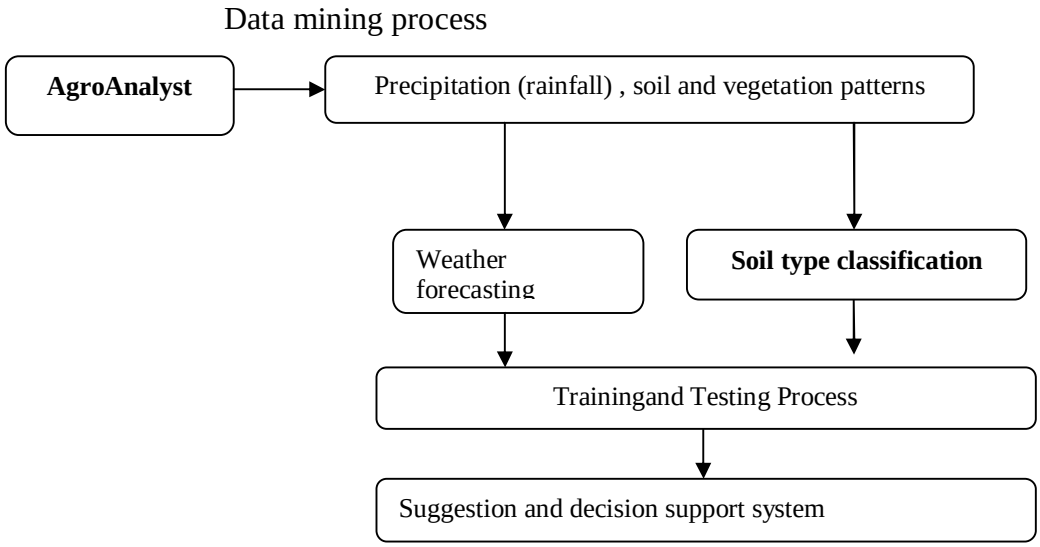

Fig 1.0Four step process of the proposed system

C. *PROPOSED ALGORITHM USING C5.0*

The agricultural decision tree algorithm made a number of changes to improve C4.5 and C5 algorithms some of these are:

• The proposed system handles training data with missing values of attributes. So, the prediction will be more accurate and effective.
• Handling differing data typed features.
• Prediction probabilities based on the historical dataset
• Pruning the decision tree after its creation
• Improves the performance of the weak classifiers by modifying the process on it.
• Handling attributes with discrete and continuous values Let the training data be a set R= R1, R2 ... of already classified samples. Each sample Di = Rl , R2... is a vector where Rl , R2 ... represent attributes or features of the sample.

The training data is a vector V = V1, V2..., where V1, V2... represent the class to which each sample belongs to. At each node of the tree, C4.5 chooses one attribute of the data that most effectively splits data set of samples T into subsets that can be one class or the other [6]. It is the normalized information gain (difference in entropy) that results from choosing an attribute for splitting the data. The attribute factor with the highest normalized information gain is considered to make the right decision for the selected test samples.

(Enhanced C5.0)

The system develops a new decision tree algorithm using many innovative features. The current study deals the problem of combinatorial optimization problem, which is not handled in the literature. So it created a new algorithm

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

named as agri decision tree algorithm. The system performs an iterative method to improve the accuracy and efficiency. This also aims at reducing the training complexity in the classification of agricultural details.

The EC5.0 exploits frequent items and rule set to build accurate EC5training method. Consider a labeled structured data set D, a frequency and the rule threshold, and a test case T. EC5exploits the novel approach to estimate, for each class in D. EC5is a novel iterative classifier. It evaluates and selects on the fly a set of eligible item sets when a new test case T has to be classified. To allow efficient pattern retrieval, EC5first performs data storage in a disk-based compact data representation. The followings are the major process included in the proposed study.[7]

- Weather prediction and soil feature extraction
- Training phase
- Test Phase

In this section, the EC5is thoroughly described. This section presents the training phase of EC5. The test phase of EC5 is based on the basic statistical evaluation process.

III. EXPERIMENTS RESULTS

In this research, the experiment uses the real time and as well as synthetic data sets for experiments. In particular, initially the data set has been collected from different sources. The Coimbatore region dataset are collected from the TNAU, its website etc., the dataset includes location; soil details and appropriate decision are collected from literature. The system can have n number of tuples for experiments. [8]

Dataset 1: (a) Soil Feature dataset with several properties for three types of soils named as sand, silt and clay.

Property/Behavior	Sand	Silt	Clay
Surface area to volume ratio	Low	Medium	High
Water-holding capacity	Low	Medium to high	High
Ability to store plant nutrients	Poor	Medium to high	High
Nutrient supplying capacity	Low	Medium to high	High
Aeration	Good	Medium	Poor
Internal drainage	High	Slow to medium	Very slow
Organic matter levels	Low	Medium to high	High to medium
Compactability	Low	Medium	High
Suceptibility to wind erosion	Moderate	High	Low
Suceptibility to water erosion	Low	High	Low if aggregated, high if not
Sealing of ponds and dams	Poor	Poor	Good
Pollutant leaching	Poor	Medium	Good

(After Brady and Weil, 2008)

In this table explain about soil property. I have selected 3 soil types in that checking the behavior and performance.

Dataset 1: (b) Soil Feature dataset with several properties for three types of soils named as sand, loam and clay.

Texture	component	intake_rate	water_retention	drainage	class
coarse	sand	very_high	very_low	low_erosion	sandy
coarse	loamy_sand	very_high	very_low	low_erosion	sandy
coarse	loamy_sand	high	low	good_drainage	sandy
moderate	sandy_loam	Moderate_high	Moderate_low	low_erosion	loamy
moderate	fine_loam	Moderate_high	Moderate_low	low_erosion	loamy
moderate	fine_loam	Moderate_high	Moderate_low	good_drainage	loamy
medium	very_fine_loam	medium	Moderate_high	moderate_draina...	loamy
medium	loam	medium	Moderate_high	moderate_draina...	loamy
medium	silty_loam	medium	Moderate_high	moderate_draina...	loamy
medium	silt	medium	Moderate_high	moderate_draina...	loamy
moderate_fine	clay_loam	moderate_low	high	?	loamy
moderate_fine	sandy_clay_loam	moderate_low	high	?	loamy
moderate_fine	sity_clay_loam	moderate_low	high	?	loamy
fine	sity_ sand	low	high	drainage	clay
fine	sity_ sity_clay	low	high	severe_erosion	clay
fine	clay	low	high	severe_erosion	clay

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

After the data set is generated, and given the number of m classes, each tuple from the synthetic dynamic database D is assigned to database Si chosen uniformly. Clearly, all datasets are formatted in a same manner. In particular, a the dataset are converted into desired type for different type of implementation sample set of the underlying data set, and the sample sets are mutually disjoint.[9]

Dataset 2: (a) Weather details which are crawled from web using web API. This includes the current and forecasted weather for the given location.

id	From_time	rainval	rain_det	temp_min	temp_...	humidity	cloudy	cloud_det	wind_sp...	wind_percen	WDate
Thondamuthur	2016-10-05T15:00:00	800	clear sky	21.66	21.66	89	clear sky	8	3.68	Gentle Breeze	05/10/2016
Thondamuthur	2016-10-05T18:00:00	803	broken clo...	20.58	20.58	97	broken clouds	64	3.03	Light breeze	05/10/2016
Thondamuthur	2016-10-05T21:00:00	500	light rain	20.63	20.63	95	scattered clo...	44	2.31	Light breeze	05/10/2016
Thondamuthur	2016-10-06T00:00:00	500	light rain	23.4	23.4	80	few clouds	24	2.79	Light breeze	06/10/2016
Thondamuthur	2016-10-06T03:00:00	801	few clouds	29.26	29.26	57	few clouds	20	3.96	Gentle Breeze	06/10/2016
Thondamuthur	2016-10-06T06:00:00	802	scattered ...	29.52	29.52	58	scattered clo...	36	5.48	Moderate br...	06/10/2016
Thondamuthur	2016-10-06T09:00:00	500	light rain	25.13	25.13	78	broken clouds	56	4.91	Gentle Breeze	06/10/2016
Thondamuthur	2016-10-06T12:00:00	500	light rain	22.48	22.48	91	broken clouds	56	4.17	Gentle Breeze	06/10/2016
Thondamuthur	2016-10-06T15:00:00	500	light rain	21.65	21.65	93	broken clouds	76	3.61	Gentle Breeze	06/10/2016
Thondamuthur	2016-10-06T18:00:00	500	light rain	21.55	21.55	95	broken clouds	80	3.36	Gentle Breeze	06/10/2016
Thondamuthur	2016-10-06T21:00:00	500	light rain	21.49	21.49	96	broken clouds	68	3.42	Gentle Breeze	06/10/2016
Thondamuthur	2016-10-07T00:00:00	500	light rain	23.54	23.54	84	scattered clo...	32	3.56	Gentle Breeze	07/10/2016
Thondamuthur	2016-10-07T03:00:00	800	clear sky	29.15	29.15	59	clear sky	8	4.23	Gentle Breeze	07/10/2016
Thondamuthur	2016-10-07T06:00:00	800	clear sky	30.55	30.55	52	clear sky	0	5.25	Gentle Breeze	07/10/2016
Thondamuthur	2016-10-07T09:00:00	800	clear sky	27.68	27.68	57	clear sky	0	5.26	Gentle Breeze	07/10/2016
Thondamuthur	2016-10-07T12:00:00	800	clear sky	23.38	23.38	74	clear sky	8	4.17	Gentle Breeze	07/10/2016
Thondamuthur	2016-10-07T15:00:00	802	scattered ...	21.53	21.53	89	scattered clo...	44	3.77	Gentle Breeze	07/10/2016

This includes the date time, locaiton, rain details, humididty , wind and cloud details.

[10]This evaluates the efficiency of the algorithms, in terms of time consumption against dimensionality d, number of traversal of tree, and tree generation threshold q under two distributions of different locations. This also evaluates the progressiveness of the methods under different datasets.

This section evaluates the proposed AgroAnalyst with EC5.0 data framework in terms of both tree traversal overhead and accuracy. We applied AgroAnalyst on sample agri data namely, Coimbatore north, PN palayam and Thondamuthur series on final set of experiments.

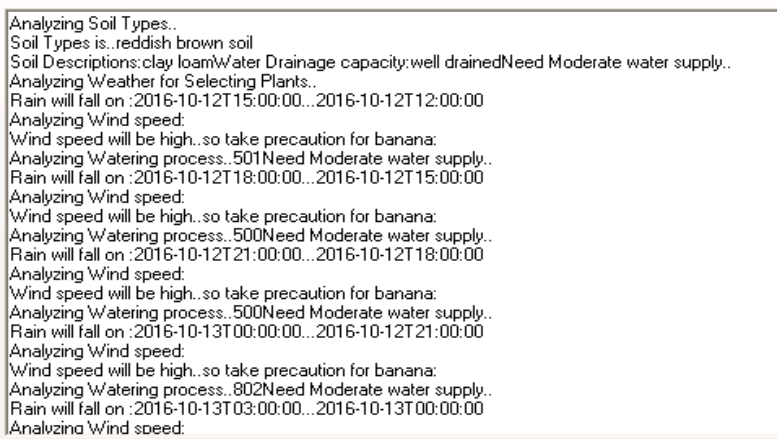

Fig:2.0 Sample output of the proposed system

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

The above figure represents the suggestions at every iteration based on the current weather and soil details. This result show soil type, soil description, and wind speed, rain fall for suggesting plant.

IV. CONCLUSION AND FUTURE WORK

Agriculture processing is tough to learn, when the user doesn't know anything about the agricultural activities. This toughness is due to the high dimensional and dynamic data. The farming and other agricultural activities are varying according to the location, soil and weather details. This work presented an overview of the agricultural data required for Agricultural DSSs for decision making and provides an interactive GUI based tool to analyse the soil, weather and appropriate plants for that. The decisions related to the Agricultural activities are highlighted.
When handling the weather and soil details, the inputs are dynamic and not common for all area. This creates a decision making problem. This study handles Coimbatore district data for creating an agroanalyst tool. This includes the enhanced C5.0 algorithm for fast decision support process. This finds the appropriate solutions and decisions based on the given attributes and values. The experiments are conducted with various conditions and factors to evaluate the output of the proposed system. And the experiments are carried out using the agroanalyst is implemented in C#.net. However, further research still requires improving accuracy and precision of data, for example, developing new sensors devices and new systems such as wireless sensor network. In addition, training using sensor devices properly for farmers is another research area to assist in the collection of valuable data for decision making

REFERENCES

1. Piatestsky-Shapiro, Gregory. Advance in knowledge discovery and data mining. Eds. Usama M. Fayyad,PadhraicSymth, and RamasamyUthurusamy, Vol.21,Menlo Park: AAAI press, 1996.
2. Han, Jiawei, Jian Pei, and MichelineKamber. *Data mining: concepts and techniques*. Elsevier, 2011.
3. http://www.tradingeconomics.com/india/gdp-from-agriculture.
4. http://www.tnau.ac.in/dr/zonepdf/WesternZone.pdf.
5. http://tnhorticulture.tn.gov.in/horti/coimbatore

6.Bujlow, Tomasz, TahirRiaz, and Jens Myrup Pedersen. "A method for classification of network traffic based on C5.0 Machine Learning Algorithm."*Computing, Networking and Communications (ICNC), 2012 International Conference on*.IEEE, 2012.

7.Kesavaraj, G., and S. Sukumaran. "A study on classification techniques in data mining." Computing, Communications and Networking Technologies (ICCCNT), 2013 Fourth International Conference on.IEEE, 2013.

8. Fathima, G. Nasrin, and R. Geetha. "Agriculture crop pattern using data mining techniques." *International Journal of Advanced Research in Computer Science and Software Engineering* 4.5 (2014): 781-6.

9. Mucherino, Antonio, PetraqPapajorgji, and Panos M. Pardalos. "A survey of data mining techniques applied to agriculture." *Operational Research* 9.2 (2009): 121-140.

10. Wheeler, Timothy R., et al. "Temperature variability and the yield of annual crops." *Agriculture, Ecosystems & Environment* 82.1 (2000): 159-167.