AGRICULTURAL ANALYSIS FOR NEXT GENERATION HIGH TECH FARMING IN DATA MINING

[1]S.Kavitha, [2]D.Geetha, [3]M.Gomathi, [4]R.Suresh Kumar

[1]M.Phil Research Scholar, [2]HOD, [3]Teacher, [4]Assistant Professor
Department of computer Science

***ABSTRACT*: Recent developments in Information Technology for agriculture field have become an interesting research area to predict the crop yield [1]. In today's world, the amount of information stored has been enormously increasing day by day which is generally in the unstructured form and cannot be used for any processing to extract useful information using mining technique [2]. This paper presents a brief analysis of data mining methods and agriculture techniques, farm types, soil types, prediction using Multiple Linear Regression (MLR) technique for the selected region. This work mainly focuses on analyzing the agricultural analysis of organic farming and inorganic farming, time cultivation of the plant, profit and loss of the data and analyzes the real estate business land in a specific area and comparison of irrigated and unirrigated land. It concentrates organic, inorganic and real estate data sets from which the prediction in agriculture will be achieved. The purpose is to estimate difference in efficiency and prediction between organic and inorganic farming. This work aims at finding suitable data models that achieve a high accuracy and a high generality in terms of yield prediction capabilities.**

***Keywords: Data mining, Agriculture Techniques, Farm Types, Soli Types*.**

1,INTRODUCTION

Agricultural researchers and farmers deploy sensors at their remote agricultural-fields to obtain the data of temperature, humidity, soil moisture and so on. Automatic collection of those data greatly helps their analytical works. Currently, they are relying on network providers, for example, cellular phone network to achieve such automatic collection from their remote sites, which is not feasible for most of the farmers due to operational cost. The review and meta-analysis of yield data comparing organic and conventional agriculture showed that currently organic yields of individual crops are on average 80% of conventional yields [1]. The analysis of 362 datasets also showed a high variation of the yield gap of organic agriculture (standard deviation 21%). Some of this variation seems systematic. E.g. soybean, some other pulses, rice and corn score higher than 80% and wheat, barley and potato scoring lower than 80%. Most regions have relative yields fairly close to the overall average.

This paper is organized as follows:

- Section 2 explains data mining
- Section 3 explains various techniques for agriculture
- Section 4 explains types of farms
- Section 5 explains types of soil
- Section 6 explains scope
- Section 7 comparison table

2. Data mining

Data mining has attracted a great attention in the information industry and in society as a whole in recent years, due to wide availability of huge amount of data and the imminent need for turning such data into useful information and knowledge [3]. The information and knowledge gained can be used for application ranging from market analysis, fraud detection, production control, disaster management and science exploration. Data mining can be viewed as a result of the natural evolution of information technology. The database system industry has witnessed an evolutionary path in the development of various functionalities: data collection and database creation, database management (including data storage and retrieval, database transaction processing and advance data analysis). Knowledge discovery as a process consists of an iterative sequence of following steps:

1. Data cleaning, that is to remove noise and inconsistent data.
2. Data integration, that is, where multiple data sources are combined.
3. Data selection, that is, where data relevant to the analysis task are retrieved from the database

4. Data transformation, that is, where data are transformed or consolidated into forms

5. Data mining, that is, an essential process where intelligent methods are applied in order to extract the data patterns.

6. Knowledge presentation, that is, where visualization and knowledge representation techniques are used to present the mined knowledge to the user

3.Various technique for agriculture

These are the following techniques used in data mining:

3.1. Genetic Algorithm

In the field of weather forecasting, a **genetic algorithm** is a search method that mimics the process of natural selection . Genetic algorithms belong to the higher class of evolutionary algorithms, which generate solutions to optimization problems using techniques inspired by natural growth , such as inheritance, mutation, selection and crossover. Genetic Algorithms are adaptive method search algorithm based on the consideration of natural selection and genetics. As such they represent an intelligent making use of a random search used to solve optimization problems. Although randomized, GAs are by no means random, alternatively they use historical information to direct the search into the region of better performance within the search space. The common techniques of the GAs are designed to simulate processes in natural systems needed for evolution mostly those follow the principles first laid down by Charles Darwin of "survival of the fittest." Since in nature, dispute among individuals for scanty resources results in the fittest individuals dominating over the weaker ones.

3.2 Artificial Neural Network (ANN)

An Artificial Neural Network is a data processing paradigm that is inspired by the method biological nervous systems, such as the brain, process data. It is composed of a huge number of highly interconnected processing elements (neurons) working in simultaneously to solve specific problems. An ANN is configured for a fixed application, such as data classification, through a learning undertaking. The artificial neuron is a data processing unit that is fundamental to the action of a neural network [5].

3.3. Nearest neighbor

Nearest neighbor search, also called as proximity search, similarity search or closer point search, is a methodology problem for discovering the closest points. Closeness is typically shows in terms of a variation function: the less similar the objects, the massive function values. Formally, the nearest-neighbor search problem is explained under as : given a set S of points in a memory M and a query point $q \in M$, find the closest point in S to q. Donald Knuth in *The Art of Computer Programming* called it the stake-office problem, mention to an application of assigning to an accommodation the closer post office. A direct generalization of this issue is a k-NN search, where we need to find the k closest points. Most generally M is a metric space and variation is expressed as a distance metric, which is symmetric and satisfies the triangle inequality. Even more familiar , M is taken to be the d-dimensional vector space where variation is scalable using the Euclidean distance, Manhattan distance or other interval metric. However, the variation function can be arbitrary.

3.4. Rule Induction

Rule Induction reasoning is a logical process in which multiple premises, all believed true are found true most of the time, are merged to secure a specific end. Inductive reasoning is often used in applications that involve prediction, forecasting, or behavior. A meteorologist will tell you that in the United States, most tornadoes rotate counter clockwise, but not all of them do. Accordingly, the final summary is probably true, but not necessarily true. Inductive reasoning is, far from deductive reasoning, not logically rigorous. Imperfection can exist and inaccurate ending can occur, however rare; in deductive reasoning the final summary are mathematically certain. Inductive reasoning is sometimes added with mathematical selection, an entirely different process [6].

3.5. Memory based reasoning

In MBR systems, reasoning is directly by using databases, so there are many advantages [1]. No knowledge asset from experts is necessary. Inserting and Eliminating knowledge, and explanation of answers is easily done. Systems can be quickly prototyped. This is in striking contrast to rule-based reasoning systems, which have difficulties with knowledge asset, with maintenance costs and enlargement time. In addition, the algorithm used for MBR is inherently data-parallel, and so can be executed on a similar machine, making full use of its powerful capabilities [7].

4.Types of farms

4.1 SUBSISTENCE AND MODERN (INTENSIVE) FARMING

Subsistence farming and Intensive farming are two ways of cultivation and differ in their objectives. Farming dates back in 8000 BC, it used to be one of the primary way of life in every country. It is the main source for provision. However, as centuries unfold, various types of farming have been made by man. Some of these are Subsistence farming and Intensive farming[1].

Subsistence Farming

Subsistence farming is used as a primary way for a family or a community to have food served on their table, the whole year round. It is when they just plant and cultivate crops for their own consumption based on their own calculation of the needed produce for the whole month or year. Farmers make sure that they have enough to last their family and no profit is intended for this.

4.2Modern (Intensive) Farming

Modern (Intensive) farming is for mass production of crops that can supply enough for a whole lot of consumers. It utilizes large land area with big investments on using labor, fertilizers and pesticides. The main reason for this type of farming is to gain profit. Since it is used for commercial production, this makes use of the latest machinery and technology to further enhance its output.

4.3 ALTERNATIVE FARMING

When we make a review about agricultural history, we can call agriculture as an occupational field which people have been dealing with since the history of mankind when the first human is created in a profession. According to the date of the existing conjectures, people first felt the need of food by hunting as a hunting society, and then the society discovered farming by developing as a collecting society.

Since the old farming style till now, people have found a lot of methods that will ease agriculture and upon the production of steamed and later on the oil machines more people have started to deal with farming. So farming has become a kind of work that can easily be done. There are now opportunities for alternative farming by the help of new farming methods.

4.4Alternative Farming Methods

Organic Farming

Organic farming is an agricultural method that includes, plant rotation, green manure, compost, biological harmful control, and it based on the mechanic tillage to ensure soil productivity, and also it refuses or restricts synthetic fertilizer, pesticide, animal feed additives and genetically modified organism.

The Importance of Organic Farming First of all Organic farming is agriculture within the framework of certain rules. This is particularly the soil, including water, air, and animals do not harm other creatures living in the environment and protecting nature.

5. Types of Soil

Soil classification deals with the systematic categorization of soils based on distinguishing characteristics as well as criteria that dictate choices in use. Soil classification is a dynamic subject, from the structure of the system itself, to the definitions of classes, and finally in the application in the field. Soil classification can be approached from the perspective of soil as a material and soil as a resource, the dataset was collected from the various resources [8].

The soil types of a particular area play critical role is determining the fertility status and cropping pattern. Red Calcareous soil, Black soil and Red non-calcareous soil are major soil types. Percentage distribution of Red Calcareous soil is high as compared to other soil types. The soil is predominantly black soil which is suitable for cotton crop. The other important causes for decline in crop production response to the application of inputs and technology is the gradual degradation of soil, the key factor for sustaining agriculture. The imbalanced fertilizer consumption, without taking into account the soil needs and soil health is proving counterproductive.

Soil types

Major Soils	Area (‘000 ha)	Percent (%) of total
Deep black soil	87.6	18.1
Deep red soils	86.5	17.8
Moderately deep black soils	74.6	15.4
Moderately deep red soils	35.9	17.4
Moderately shallow black soils	6.2	1.3
Moderately shallow red soils	65.7	13.6
Shallow black soils	2.9	0.6
Shallow red soils	59.4	12.3
Very deep black soils	22.9	4.7
Very deep red soils	7.5	1.6
Very shallow black soils	4.6	1.0
Very shallow red soils	1.8	0.4

6. Scope

Making agriculture sustainable and resilient to the ongoing change in climate and social structure is a major challenge for the scientists and researchers across the globe. Agricultural system demands transition and a multidisciplinary approach. Intelligent and precision agricultural approaches were given due importance for increasing production and productivity from the very same limited resources. The approach needs information from various sources and efficient use of them in relevant field. This need lead to growing interest in knowledge discovery from vast piles of data generated out of various research and survey works. The emergence of Data Mining techniques revolutionized the field of information generation and pattern recognition. Though Data Mining is an emerging science, it finds a wide application in agriculture and allied sectors, and has a wide future prospect Agriculture,

Agriculture is experiencing a transition stage driven by population pressure and climate change. More production and productivity are being expected from
Limited resources. New intensive research is being done to explore ways to increase production with optimum use of resources maintaining sustainability. This lead to the use of modern sophisticated computer assisted technologies in agricultural research. Due to widespread use of computer and affordable storage facilities, there is an enormous wealth of data embedded in huge databases of different agric-allied enterprises.

A recent development in Information Technology for agriculture field has become an interesting research area to predict the crop yield In today's world, the amount of stored information has been enormously increasing day by day which is generally in the unstructured form and cannot be used for any processing to extract useful information using mining technique. This paper presents a brief analysis of crop yield prediction using Multiple Linear Regression (MLR) technique for the selected region. This project is mainly focused on analyzing the agriculture analysis of organic forming and inorganic forming ,time cultivation of the plant, profit and loss of the data and analyze the real estate business land in a specific area. The research on comparison for irrigated and unirrigated land. It concentrates organic, inorganic and real estate data sets from which the prediction in agriculture will be achieved. The purpose is to estimate difference in efficiency and prediction between organic and inorganic forming.

7. Comparision table

Following tabular form is used to find various weather prediction techniques and its advantage and disadvantage.

Author	Title	Data mining methodologies	Applications	Advantages	Disadvantages
Sanjay D. Sawaitul, Prof. K.P. Wagh, Dr. P.N. Chatur	Classification and Prediction of Future Weather by using Back Propagation Algorithm- An Approach	Neural Networks	Focuses on weather forecasts	Good prediction accuracy	Small application of cart. Only work in nominal variable
V. K. Somvanshi, et al.,	Modeling and prediction of rainfall using artificial neural network and ARIMA techniques	Neural Networks	Prediction in rainfall	Rainfall prediction 83.42% accuracy.	Only applicable in rainfall prediction.
I. Jagielska, C. Mattehews, T. Whitfort	An investigation into the application of neural networks, fuzzy logic, genetic algorithms, and rough sets to automated knowledge acquisition for classification problems	K-means	Classifying soil in combination with GPS	Accurate results in advance prediction of agricultural	It's focused on only observed parameters information.

S.Veenadhari, Dr. Bharat Misra, Dr. CD Singh	Data mining Techniques for Predicting Crop Productivity	Decision Tree Analysis	Influence of climatic factors on major kharif and rabi crops production	Simple to understand. Large amount of data can be analyzed.	Poor accuracy. It suffers the over fitting problem.
K. Verheyen, D. Adriaens, M. Hermy, and S. Deckers	Adriaens, M. Hermy, and S. Deckers High resolution continuous soil classification using morphological	Fuzzy set	Yield Prediction in agriculture	To minimizes the losses and damages. Results are reached 95%.	Prediction is limited by unknown conditions.
Veenadhari, S.	Crop productivity mapping based on decision tree and Bayesian classification	K-nearest Neighbor	Scale back procedure burden of k-nearest neighbor algorithms	To have a better result	Small size of data available in training testing.

V. CONCLUSION

Agriculture is the most significant application area particularly in the developing countries like India. Use of information technology in agriculture can change the situation of decision making and farmers can yield in better way. Project technique plays a crucial role for decision making on several issues related to agriculture field. The systems must be compared at rotation level using a unit (e.g. grain equivalents) that allows for comparing farming systems with different crop compositions. It discusses about the role of data mining in the agriculture field and their related work by several authors in context to agriculture domain. It also discusses on different data mining applications in solving the different agricultural problems. This paper integrates the work of various authors in one place so it is useful for researchers to get information of current scenario of data mining techniques and applications in context to agriculture field.

REFERENCES:

[1] Data mining techniques for agriculture: A review International journal on recent and innovation trends in computing and communication (volume: 2issue:8) IJRITCC| August 2014, available

[2] International Journal of Advanced Engineering Research and Science (IJAERS)

[3]Application and Scope of Data Mining in Agriculture [Vol-3, Issue-7, July- 2016] ISSN: 2349-6495(P) | 2456-1908(O) www.ijaers.com

[4] International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) A Survey on Data Mining Techniques in Agriculture Vol. 3, Issue 2, February 2015

[5] C. Romero, S. Ventura, E. Garcia, "Data mining in course management systems: Module case study and tutorial", Computers & Education, Vol. 51, No. 1, pp. 368-384, 2008

[6] Rushing J.R., Ramachandran U, Nair S., Graves R., Welch,LinA.,2005,"A Data Mining Toolkit for Scientists and Engineers", Computers & Geo sciences, 31,607-618

[7]Developing innovative applications in agriculture using data mining. Department of Computer Science University of Waikato Hami