IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 40-42

Integrating Crop and Resource Management technologies to Enhanced Productivity, Profitability, and Sustainability of the Agricultural system"

Prof .Surekha R. Janrao[1], Prof. Nilesh Kulal[2]
[1](Department of Computer Engineering, Terna Engineering College, Nerul,Navi Mumbai)
[2](Department of Computer Engineering, Terna Engineering College, Nerul,Navi Mumbai)

Abstract: *In our India more than 70% people are dependent on agriculture profession. Data mining represent an important area in agriculture sector for knowledge discovery which can be helpful for the farmers in their profession. Situation of decision making for agricultural people has been changed due to more advanced techniques in information technology. Many past governments do so many efforts for improving the standard for agricultural people especially in rural areas. Government also launched one web portal like "ONE STOP SHOP FOR FARMERS". In this portal various information has been given for the agricultural people who will be helpful for their profession. This paper proposes improved machine learning approach which can be useful and additional information for agricultural people as well as farmer's government portal .This is the most popular and powerful approach in Data Mining. This machine learning technique is used to design classification model. This model will be used more efficiently and make a prediction for crop management which will be helpful for the farmers.*

Keywords: *Data mining, Information Gain, Classifier, KDD, Knowledge database*

I. Introduction

India is generally an agricultural country. Agriculture is the single most important contributor to the Indian economy. Agriculture crop production depends on the season, biological, and economic cause. The prognosticating of agricultural yield is a challenging and desirable task for every nation. Agriculture prediction is a very important problem in country. Any farmer is interested in knowing how much yield he is about to expect. In the past, agriculture prediction was performed by considering farmer's experience on particular field and crop. Nowadays, Farmers are struggling to produce the yield because of unpredictable climatic changes and drastically reduce in water resource so we are creating a classification model for wheat crop management based on the agricultural data. This data could be gathered, stored and analyzed for useful information. It is used to promote new advanced methods and approaches such as data mining that can give the information of the previous results to the wheat crop yield estimation. This module provides the decision to the farmer whether he will produce wheat crop or not based on the values of attributes. This decision of the wheat crop management improves the value and gain of the farmer**.**

II. Existing System

The previous analysis system contains only agricultural information which will affiliate farmer for data availability on a customer farmer portal with crops details and its users and researches can get online information about the crops. Previous analysis system prediction is concerned only about the crop disease identification. For agricultural sector in Indian Society it is very difficult for the Farmers society to provide the detailed information about the resources and crop management. And directly meet the Customers without the intermediate party or organization. Various old methods are available which is being implemented till now causes a big loss for farmers society and a minor gain for Indian economy [1].

III. Proposed System

The proposed system helps farmers in all manners that are crop analysis and management understanding it more clearly. Initially the statistical classification model i.e. decision tree is applied on existing data and it will be again compare with Bayes classification method. Proposed system will make the decision based on the classification model that is designed from the training data set and that model further used the new data set to predict the future values.

- It provides an easy useful and secure way for crop management and extraction of large data sets.
- Number of crops availability is huge in numbers so managing agricultural data analysis become more manageable and easy.

- Farmer facing problems for their crop management so this system will improve their profitability and sustainability as they will get direct solution to their problem by implementing this model[2],[3].

IV. ALGORITHM and TECHNIQUES

Many techniques have been developed for learning rules and relationships automatically from diverse data sets, to simplify the often tedious and error-prone process of acquiring knowledge from empirical data. Decision tree is one of learning algorithm which possesses certain advantages that make it suitable for discovering the classification rule for data mining applications. Normally Decision trees widely used learning method and do not require any prior knowledge of data distribution, works well on noisy data .It has been applied to classify Wheat disease based on the symptoms. Expert systems have been used in agriculture since the early 1980s. This proposal explores what Classification model can do in the agricultural domain. Decision trees have become one of the most powerful and popular approaches in knowledge discovery and data mining, the science and technology of exploring large and complex bodies of data in order to discover useful patterns. The area is of great importance because it enables modeling and knowledge extraction from the abundance of data available. The construction of decision tree classifiers does not require any domain Knowledge or parameter setting, and therefore is appropriate for exploratory Knowledge discovery. The Decision tree can handle high dimensional agricultural data. The learning and classification steps of decision trees induction are simple and fast. The transfer of experts from consultants and scientists to agriculturists extends workers and farmers represent a bottleneck for the development of agriculture on the national. The term Knowledge Discovery in Databases or KDD for short, refers to the broad process of finding knowledge in data, and emphasizes the "high-level" application of particular data mining methods. It is of interest to researchers in machine learning, pattern recognition, databases, statistics, artificial intelligence, knowledge acquisition for expert systems, and data visualization. The unifying goal of the KDD process is to extract knowledge from data in the context of large databases. Many machine learning schemes can work with either symbolic or numeric data, or a combination of both, and attempt to discover relationships in the data that have not yet been hypothesized. Once a Relationship has been discovered, further statistical analysis can be performed. Sometimes, both fields work independently towards the same goal, as in the case of ID3 a machine learning scheme, and CART, standing for "classification and regression trees," a statistical scheme. These methods both induce decision trees using essentially the same technique. Machine learning researchers also incorporate statistics into learning schemes directly, as in the case of the Bayesian classification system AUTO CLASS. **C4.5** performs top down induction of Decision trees from a set of examples. Typically, a training set will be specified by the user. The root of the tree specifies an attribute to be selected and tested first, and the subordinate nodes dictate tests on further attributes. The leaves are marked to show the classification of the object they represent. An information-theoretic heuristic is used to determine which attribute should be tested at each node, and the attribute that minimizes the entropy of the decision is chosen. C4.5 is a well-developed piece of software that derives from the earlier ID3 scheme.

In previous analysis system prediction is concerned only about the wheat disease identification crop yield. Farmers are usually enabled to recall exactly when and for how much they have made investments and how long those investments were to be utilized. 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. The data analysis of this project which used different machine learning technique are are very fast, accurate, reliable, storage and saves a lot of time of agricultural scientists. The ID3 algorithm is a decision tree building algorithm which determines classification of objects by testing values of their properties. It builds tree in top down fashion, starting from set of objects and specification of properties. At each node of tree, the properties tested and the result is used to partition data object set. The information theoretic heuristic is used to produce shallower trees by deciding an order in which to select attributes. The first stage in applying the information theoretic heuristic is to calculate the proportions of positive and negative training cases that are currently available at a node. In the case of the root node this is all the cases in the training set. A value known as the information needed for the node is calculated using the following formula where p is the proportion of positive cases and q is the proportion of negative cases at the node[4],[5].

The decision tree classifier applied on the dataset uses three different splitting criteria namely
(i) Information Gain
ii) Gini Index
iii) Gini Ratio

The basic algorithm of ID3Examples *S*, each of which is described by number of attributes along with the class attribute C, the basic pseudo code for the ID3 algorithm is if (all examples in *S* belong to class *C*) then make leaf labeled *C* Else select the "most informative" attribute A Partition S according to A's values (v1... vn)

Recursively construct sub-trees T1, T2... Tn for each subset of *S*.ID3 uses a statistical property, called information gain measure, to select among the candidates attributes at each step while growing the tree. To define the concept of information gain measure, it uses a measure commonly used in information theory, called entropy. The entropy is calculated by,

$$Entropy(S) = \sum_{i=1}^{c} -Pi \log_2(Pi)$$

Where S is a set, consisting of s data samples, Pi is the portion of S belonging to the class i. Notice that the entropy is 0 when all members of S belong to the same class and the entropy is 1 when the collection contains an equal number of positive and negative examples. If the collection contains unequal numbers of positive and negative examples, the entropy is between 0 and 1. In all calculations involving entropy, the outcome of all calculations involving entropy, the outcome of (0 log2 0) is defined to be 0. With the Information gain measure, given entropy as a measure of the impurity in a collection of training examples, a measure of effectiveness of an attribute in classifying the training data can be defined. This measure is called information gain and is the expected reduction in entropy caused by partitioning the examples according to this attribute. More precisely, the information gain is Gain(S, A) of an attribute A, relative to a collection of examples S [6], [7].

Attributes for Wheat Crop Management are as follows:
Soil Types [Alluvial, Black, Laterite, Mountain, Red]
Climate [Cool, Warm, Moist, hot]
Agricultural Land [Small, Medium, Large]
Rainfall [Low, Medium, High]
Water Resources [Yes, No]
Temperature [Min, Max]
Class Label Attribute is, Production__Wheat [Yes, No]

From agriculture government of India web site we will collect the data set for crops and related climatic conditions as stated above and after preprocessing transformed data will be given to classifier model and further that model do the analysis of data and provide the prediction in terms of the decision for crop yield production.

V. Conclusion

This paper is a solution to provide an overview of some previous researches and studies done in the direction of applying data mining and specifically mining techniques in the agricultural domain and the profitability of farmers work enhancement. We have also tried to evaluate the current status and possible future trends in this area. Agriculture or farming forms the backbone of any country economy, since a large population lives in rural areas and is directly or indirectly dependent on agriculture for a living. The large amount of data generated and stored can be used in the process. The inclusion is simple but effective techniques will help in development of the agriculture and industrial fields. This work performs the minimum statistics on agricultural data more efficiently and easily [8]. It is an initial proposal to show that Online Agricultural Product System is feasible and convenient. Major benefit of this type of information system to Indian agricultural sector is when it became operational as planters, importers, exporters and growing the economy as well for farmers ,customers and Country for sure.

References

[1]. Hetal Patel, Dharmendra Patel, ”A Brief survey of Data Mining Techniques Applied to Agricultural Data”International Journal of Computer Applications,Volume , 9, June 2014.
[2]. Sanket Salvi, Pramod Jain S, Sanjay H., Harshita T, M. Farhana, Naveen Jain, Suhas M,“Cloud Based Data Analysis and Monitoring of Smart Multi-level Irrigation System UsingIoT",International conference on I-SMAC (I-SMAC 2017).
[3]. M.M.Wakhare,a.G.Choure,“Design of framework for agriculture web service ”, International Journal of Advances in Electronics and Computer Science, ISSN: 2393-2835 Volume-4, Issue-1, Jan.-2017
[4]. Agraj Aher, Janhavi Kasar, Palasha Ahuja , Varsha Jadhav,“Smart Agriculture using Clustering and IOT”, International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 03 | Mar-2018 p-ISSN: 2395-007
[5]. Olakunle Elijah, Tharek Abdul RahmanIgbafe Orikumhi, Chee Yen Leow ,“An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges ”, IEEE INTERNET OF THINGS JOURNAL, 2327-4662 (c) 2018 IEEE.
[6]. Prof. Chandrakanth. Biradar1, Chatura S Nigudgi,” An Statistical Based Agriculture Data Analysis” International Journal of Emerging Technology and Advanced Engineering, Volume 2, Issue 9, September 2012.
[7]. Farah Khan, Dr. Divakar Singh,”Association Rule Mining in the field of Agriculture: A Survey ”, International Journal of Scientific and Research Publications, Volume 4, Issue 7, July 2014
[8]. Abdullah, A., Brobst, S., M.Umer M. "The case for an agri data ware house: Enabling analytical exploration of integrated agricultural data". Proc. of IASTED International Conference on Databases and Applications. Austria. Feb 2004.