A REVIEW ON CROP YIELD PREDICTION

[*]Nandini Kakran
[**]Dr. K.P. Jayant

Abstract

The Various climatic situation are considered that effect the local weather conditions in various parts of the world. The weather condition for the particular area have a direct effect on crop yield. The lot of research work have been done to exploring the connections between climate and crop yield .There are some conditions such as typhoons, floods and droughts effect the crop production. The growth yield prediction is the major problem that must be solved on the base of available data. Data Mining technique can be used for this purpose. There are many techniques to be used for eliminating the future crop production. This paper have a brief discussion on crop yield prediction using Multiple Linear regression (MLR) technique and Density based clustering technique.

Keywords: Crop Production Prediction, Density based Clustering, Multiple Linear Regression.

Introduction

In Indian Economy Agriculture play a important role. In India, there are multiple reasons so that the farmers does not get the expert crop yield. The crop yield is depends on the weather conditions. Rainfall conditions also influences the crop cultivation. So , the farmers of our country require a timely advice in context of future crop productivity prediction. This analysis helps the farmers to maximize the crop production.

In agriculture, crop yield prediction is the big problem. Farmers are interested in knowing, how much the crop production is about to expect. In past years, farmers performed the yield prediction by considering the previous experience on some particular crop [1]. The extraction of useful and important information from large data sets is done through Data Mining. Data mining is used in agriculture field , where the training data is to be collected from historical data and this data is used to learn how to classify future yield predictions [2]. Data mining is the process to discover new patterns from the large datasets. The paper gives details about different techniques of data mining used in agriculture yield prediction.

Data Mining tasks is divided into two categories:
First is Descriptive data mining and second is Predictive data mining.

Descriptive data mining explains the general properties of the data in the database while Predictive data mining is used to predict explicit values from known results based on the pattern determination. Prediction means using some values or fields in the database to predict unknown values of other interested variables. Predictive data mining procedure is used in crop prediction [3].

Literature Survey

[i] The paper titled "Generalized Software tools for crop are estimation and yield forecast", **Rober to Benedetti** describes the procedure that estimates the variables of interest, like land use and crop yield. The paper explains the system in different steps : Sample design, data editing and estimation[4].
[ii] The paper titled "Risk in Agriculture – A study of Crop Yield Distributions and Crop Insurance" by **Narsi Reddy Gayam** study on the basis of data

*M.Tech Scholar, Department of CSE, IIMT University, Meerut, India
** Associate Professor, Department of CSE, IIMT University, Meerut, India

collected from **INDIA** of Sugarcane and Soyabean. The paper applies the **Liffifore** method with intensive qualitative analysis to test the null hypothesis of crop yield. The results shows that crop yield are not distributed normally [5].

[iii]The paper titled "Applying data mining techniques in the field of Agriculture and Applied Sciences", have a research study on applications of data mining techniques in the agriculture field. In this paper, ID3 Algorithms, k-means, artificial neural networks and k- nearest neighbour were presented [6].

[iv] The paper titled, "Data Mining Techniques for predicting crop productivity – A review article ", review the study on applications of data mining techniques. Data mining is the new approach in crop prediction / crop forecasting. The previous information of crop yield is important. An estimate of risk and crop size helps the companies in planning the schedule of production [7].

Techniques of data mining in crop prediction

There are mainly four data mining techniques: Classification, Clustering, Association Rule Mining, and Regression.

Classification

This technique used for classifying unknown samples with the help of information provided by a set of classified samples. The set is usually a training set because the set is used in classification as trained the classification technique.

Classification technique broadly be categorized into three approaches: Supervised learning, unsupervised learning or semi- supervised learning [8].

The different classification techniques used in agriculture for crop prediction such as Bayesian Networks, Decision Tree, Nearest Neighbour Artificial Neural Network, Support Vector Machine, Fuzzy Logic, Rough Sets and Genetic Algorithms.

Clustering

The Technique clustering is focused on portioning a data set into clusters such that the points in a cluster are close to one another. Clustering technique subsets the data items in such a way that similar type of instances are grouped together, While different instances belong to different groups [9].

There are different clustering methods such as Hierarchical Method, Partitioning Method , Density Based Method , Model Based Clustering Method, Grid Based Method.

Regression

This Technique is about mapping a data item to a real- valued prediction variable through learning a function.

There are multiple applications of regression like in a forest prediction of presence of biomass, prediction of demand of a new product to the consumer [10]. In this, the model is trained to predict the continuous target. There are two methods for predictions are Linear Regression and Non-Linear Regression [9].

Association-Rule Mining

The Association-Rule Mining Technique is the efficient technique to search desired patterns from the vast amount of data in data mining. The method focuses on finding the relation among the different items in database. This Technique are used to find elements that occur repeatedly within a datasets having many independent elements. Here, we can explain it as there is a set of transactions, each transactions having a no. of items . The association rule is expression in the form X==>Y , Where X and Y are data items. In this, the transaction of the database X tends to the database contains Y [11].

Conditions to be Considered
Greenhouse Condition

Greenhouse Conditions for plants to grown had created in **Gipoonda** Greenhouses and pacific Agri-Food Research Centre in 2002. Both greenhouses are connected with gutter connected greenhouses having high walls with ventilation at roof-top. These greenhouses was a commercial greenhouse located in Delta, region, BC [12]

Crop Culture

In this, the number of stems on side shoots were adjusted during the season.

In Tomato greenhouses, the stems must be adjusted by controlling the side shoot during the changes in light conditions.

The greenhouse with name PARC side was a research centre in **Agassiz**, BC. The greenhouse had four compartments with green area to collect the data. The stem had 2.5 stems m-2 density in each compartment.

Carbon di oxide must be enriched in two compartments only. Each compartment had different irrigation and climate conditions [13].

Methodology

The main goal of crop prediction model is to crop management using some estimate of soil conditions or weather conditions.[2]
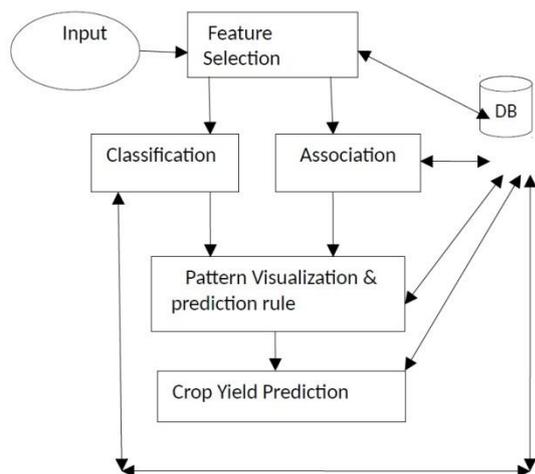
Figure: Architecture of Crop Yield Prediction

The figure shows an architecture of crop prediction with a input module, which is used for taking input from farmer, The farmer provides the region , area of land , city and economical conditions.

The city parameters are selected based on altitude, latitude and longitude . From crop knowledge base, the automatic data will be reflected.

Feature Selection Module

This module is used for selection of attribute from crop knowledge base. The crop knowledge base is the set of region-id, region- name, soil-type, water ph, rainfall, humidity, sunlight, land information, pesticides information, seed type etc.

Grouping

After attribute selection, classification and association rule is used for grouping similar contents. Then prediction rules must be applied to the cluster to get the results [8].

Conclusion

In this paper, some data mining techniques are defined to eliminate the crop production analysis with existing data. This paper presents new scope for the modern classification methodologies in the field of crop yield prediction. The data mining technique are used in the application of agriculture for crop prediction and the data amount growing per day currently available from many resources.

Reference

1. D Ramesh, B Vishnu Vardhan, "Analysis of Crop yield prediction using Data Mining Technique", IJRET, Volume 2, Issue 9, September 2013.
2. Ramesh A.Medar, Vijay.S.Rajpurohit, "A survey on Data Mining Techniques for Crop Yield Prediction", International Journal of Advance Research in Computer Science and Management Studies, Volume 2, Issue 9, September 2014.
3. M.C.S.Geetha, "A survey on Data Mining Techniques in Agriculture", International Journal of Innovative Research in Computer and Communication Engineering ", Volume 3, Issue 2, February 2015.
4. Mucherino, A., Papajorgji, P., and Pardalos, P.(2009). Data Mining in Agriculture (volume 34). Springer.
5. Narsi Reddy Gayam, "Risk in Agriculture- A study of Crop Yield Distributions and Crop Insurance", Engineering Systems Division Massachusselts Institute of Technology, June 2006.
6. YETHIRAJ. N .G, "Applying Data Mining Techniques In The Field of Agriculture and Allied Science", International Journal of Business Intelligent ISSN: 2278-2400, Volume 1 , Issue 2, December 2012.
7. S.Veenadhari, Dr. Bharat Mishra, Dr. CD Singh, "Data Mining Techniques for Predicting Crop Productivity- A review article", IJCST, March 2011.
8. Ramesh A. Medar, Vijay. S.Rajpurohit ," A Survey on Data Mining Techniques for Crop Yield Prediction", International Journal of Advance Research in Computer Science and Management Studies, Volume 2 , Issue 9 , September 2014.
9. Hetal Patel, Dharmendra Patel, "A Brief Survey of Data Mining Techniques Applied to Agricultural Data", IJCSA, Volume 95-No.9, June 2014.
10. Fayyed, U., Piatetsky- Shapiro. G., and Smyth, P. (1996). "From Data Mining to Knowledge Discovery in Databases", AI magazine.
11. Srikant, RUQ and Agrawal, R(1997, August). Mining Association Rules with Item Constraints in KDD Volume 97, pp. 67-73.
12. David L Ehret, Bernard D. Hill, Tom Helmer, Diane R. Edwards, "Neural Network Modelling of greenhouse tomato yield , growth and water use from automated crop monitoring data ", 7 march 2011.
13. Baille, M., Laury, J.C., Baille, A., 1992 "Some comparative results on evapotranspiration of greenhouse ornamental crops, using lusimeter, greenhouse H2O balance and LVDT sensors" Acta Hortic. 304, 199-208.