

Agriculture Data Analytics in Crop Yield Estimation Literature Survey.

College Name: Adhi College of Engineering and Technology

Department: Computer Science and Engineering

Team Members: Nandhu G(Team Leader)

Charanraj S

Abitha A

Indhumathi M

ABSTRACT

Analytics is the interpretation of data pattern that assist decision- making and performance improvement. Agriculture Data analytics in crop yield helps in analysing some important visualization, creating a dashboard and by going through these we will get most of the insights of Crop production in India. IBM Cognos Analytics integrates reporting, modelling, analysis, exploration, dashboards, stories, and event management so we can understand our organization's data, and make effective decisions. A dashboard helps us to monitor events or activities at a glance by providing key insights and analysis about our data on one or more pages or screens. In this project, we visualize, analyse and gain most of the insights by creating a dashboard. Keywords: IBM Cognos, Crop Production, Analytics and Agriculture

LITERATURE SURVEY

M. A. Jayaram and Netra Marad, "Fuzzy interference Systems for Crop Prediction", Journal of Intelligent Systems, 2012, 21(4), pp.363-372[1]. Prediction of crop yield is significant in order to accurately meet market requirements and proper administration of agricultural activities directed towards enhancement in yield. Several parameters such as weather, pests, biophysical and morphological features merit their consideration while determining the yield. However, these parameters are uncertain in their nature, thus making the determined amount of yield to be approximate. It is exactly here that the fuzzy logic comes into play. This paper elaborates an attempt to develop fuzzy inference systems for crop yield prediction. Physio morphological features of Sorghum were considered. A huge database (around 1000 records) of physio morphological features such as days of 50 percent flowering, dead heart percentage, plant height, panicle length, panicle weight and number of primaries and the corresponding yield were considered for the development of the model. In order to find out the sensitivity of parameters, one-to-one, two-to-one and three-to one combinations of input and output were considered. The results have clearly shown that panicle length contributes forth yield as the lone parameter with almost one-to-one matching between predicted yield and actual value while panicle length and panicle weight in combination seemed to play a

decisive role in contributing for the yield with the prediction accuracy reflected by very low RMS value.

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. Crop Yield Prediction Using Machine Learning A research group investigated the utilization of various information mining methods which will foresee rice crop yield for the data collected from the state of Maharashtra, India. A total of 27 regions of Maharashtra were selected for the assessment and the data was collected related to the principle rice crop yield influencing parameters such as different atmospheric conditions and various harvest parameters i.e Precipitation rate, minimum, average, maximum and most extreme temperature, reference trim cultivable area, evapotranspiration, and yield for the season between June to November referred as Kharif, for the years 1998 to 2002 from the open source, Indian Administration records. WEKA a Java based dialect programming for less challenging assistance with information data sets, assigning design outcomes tool was applied for dataset processing and the overall methodology of the study includes, (1) pre-processing of dataset (2) Building the prediction model utilizing WEKA and (3) Analyzing the outcomes. Cross validation study is carried out to scrutinize how a predictable information mining method will execute on an ambiguous dataset. Study applied 10-fold higher cross validation study design to assess the data subsets for screening and testing. Identified and collected information was randomly distributed into 10 sections where in one data section was used for testing while all other data sections were utilized for the preparation information. Study reported that the method applied was supportive in the precise estimation of rice crop yield for the state of Maharashtra, India. The precise quantification of the rice productivity in various climatic conditions can help farmer to understand the optimum condition for the higher rice crop yield. Simulation models based on field experiment are valuable technologies for studying and understanding crop yield gaps, but one of the critical challenge remain with these methods is scaling up of these approach to assess the data collated between different time intervals from the broader geographical regions. Satellite retrieved data have frequently been revealed to present data sets that, by itself or in grouping with other information and model designs, can precisely determine the yields of crop in agricultural lands. The yield maps developed shall provide an unique opportunity to overcome both spatial and temporal based scaling up challenges and thus improve the ideology of crop yield gaps prediction. A review was conducted to discuss the applications of remote sensing technology to determine the impact and causes of yield gaps.

Even though the example discussed by the research group demonstrates the usefulness of remote sensing in the prediction of yield gaps, but also many areas of possible application with respect to the crop yield assessment, prediction and improvement remain unexplored. Study proposed two less complicated, easily assessable methods to determine and quantify the yield gaps between various agricultural fields. First method works closely with the constructive maps representing the average crop yields, it can be used directly to access specific crop yield influencing factors for further studies whereas the second method use the remote sensing technology to retrieve the data for providing the useful information regarding the crop yield prediction and estimation .

REFERENCES

- [1] M. A. Jayaram and Netra Marad, "Fuzzy interference Systems for Crop Prediction", Journal of Intelligent Systems, 2012, 21(4), pp.363-372.
- [2] P. Vindya "Agricultural Analysis for Next Generation High Tech Farming in Data Mining" , Anna University, Trichy, Tamil Nādu, India, 5 May 2015
- [3] Dakshayini Patil, M .S, Shirdhonkar. Rice Crop Yield Prediction using Data Mining Techniques: An Overview. International Journal of Advanced Research in Computer Science and Software Engineering, 2017; 7(5):427-431.
- [4] David B. Lobell, The use of satellite data for crop yield gap analysis, Field Crops Research-143, 2013; 56–64.
- [5]<https://www.toolbox.com/hr/hr-analytics/articles/what-is-hr-analytics/>
- [6] <https://www.ibm.com/downloads/cas/WVJBQERV>