**Literature Review**

**Title :** Estimate The Crop Yield Using Data Analytics

**Domain** : Data Analytics

**Team member :** Anuraagavi M R

Aditya Venkatesh

Haripriya K

Harinivas M

**REVIEW OF LITERATURE METHODS OF CROP YIELD PREDICTION :**

At present we are at the immense need of another Green revolution to supply the food demand of growing population. With the decrease of available cultivable land globally and the decreased cultivable water resources, it is almost impossible to report higher crop yield. Agricultural based big data analytics is one approach, believed to have a significant role and positive impact on the increase of crop yield by providing the optimum condition for the plant growth and decreasing the yield gaps and the crop damage and wastage. With this aim the present paper reviews about the various advances, design models, software tools and algorithms applied in the prediction assessment and estimation of the crop yield. India is basically agriculture based country and approximately 70% our country economics is directly or indirectly related to

the agricultural crops. The principle crop which occupies the highest (60-70%) percentage of cultivable land in the Indian soil is the paddy culture and it is the major crop especially in central and south parts of the India. Rice crop cultivation plays an imperative part in sustenance security of India, contributing over 40% to

general yield generation. The enhanced yield of the rice crop depends largely on the water availability and climatic conditions. For example, low precipitation or temperature extremes can drastically diminish rice yield. Growing better strategies to foresee yield efficiency in a mixture of climatic conditions can help to

understand the role of different principle factors that influence the rice crop yield. Big data analytic methods related to the rice crop yield prediction and estimation will certainly support the farmers to understand the optimum condition of the significant factors for the rice crop yield, hence can achieve higher crop yield.

**CROP YIELD PREDICTION USING DATA ANALYTICS:**

The accurate prediction of crop yield certainly benefits the farmers in choosing the right method to reduce the crop damage and get best prices for their crops. A research group conducted a work with an objective of accurate prediction of crop yield through big data analytics to assess various crop yield influencing factors such as Area under Cultivation (AUC) interim of hectors, Annual Rainfall (AR) rates

and Food Price Index (FPI) and to develop relationship among these parameters. Regression Analysis (RA) methodology was applied to examine the selected factors and their impact on crop prediction and final yield. RA methodology is a multi-variable investigation practice which can categorize the factors into groups such

as explanatory and response variables and helps to assess their interaction to obtain a resolution.A novel method called Linear Regression (LR) is applied to analyze the

relationship between explanatory variables (AR, AUC, FPI) and the crop yield considered as response variable. The Study reported that the R2 value for the studied factors clearly indicates that crop yield is principally depends on AR. Study also reported that the other two factors (AUC and FPI) screened were also found to have significant impact after the AR. Study shall be continued to analyze the impact of for other substantial factors like Minimum Support Price (MSP), Cost Price Index (CPI), Wholesale Price Index (WPI) etc. and their relationship on the yields of different crops.

A case study was discussed on the application of selected model design to quantify the yield gaps of maize crop in the state of Nebraska (USA), and also at the different geographical locations representing the nations Argentina and Kenya at

national scale level. Different geographical locations such as Nebraska (USA), Argentina and Kenya were identified to symbolize the distinct scenarios of Agriculture based data availability and the quality of the selected variables assessed to predict and estimate the crop yield gaps. The definitive aspiration of the planned method is to afford transparent, easily accessible, reproducible and technically sound and strong guidelines for predicting the yield gaps. The proposed guidelines were also relevant for understanding and to simulate the influence of change in climate conditions and usage of cultivable land changes from national to global scales. Analyzing the yields of crop is necessary to update the policies to ensure food security. A research

The Study suggested the aim in suggesting a novel data mining method to predict the yields of crop depends on agricultural data analytics methodologies, which were progressively contrast with conventional data mining methodologies in the process of handling data and modeling designs. The study suggested that the method employed should be user-friendly, work based on progressive data responsive processing structure, supposed to utilize the existing agricultural based significant datasets and would still be used with the larger volumes of data growing at enormous rates. Nearest neighbors modeling is one such novel data mining technique which works on the results collected based on data processing

structures form the farmers and suggest a well unbiased result on the base of accuracy and prediction time in advance. Study further discussed a case study on the assessment of actual crop dataset in China from 1995-2014. Study reported that the novel model employed has publicized an improved performance and was found to be progressive in reporting prediction accuracy percentage of the compared

methodologies with conventional designs 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.