

# Agricultural Production Analysis for Smart Farming using Data Science

<sup>1</sup>P. Sakthivel, <sup>1</sup>S. S. Aravinth, <sup>2</sup>R. Radhapriya, <sup>2</sup>M. Sangeetha, <sup>2</sup>K. Pavithra

<sup>1</sup>AP/CSE, Dhirajlal Gandhi College of Technology, Salem, Tamilnadu

<sup>2</sup>UG Scholar, CSE, Dhirajlal Gandhi College of Technology, Salem, Tamilnadu

Contact Email: aravinth.cse@dgct.ac.in

***Abstract***— Smart Farming is achieved based on result analysis on datasets related to the crop variety, crop yields, crop cost. The datasets are in the. Csv file. The main purpose to increase the crop productivity. Monitoring the crop growth parameters on the basis of real time data processing through an Algorithm. Crop cultivation and the yield have been analyzed. The environmental data provided by the precision agriculture technology has been analyzed and useful information has been extracted. In agricultural, analytics is the management of structured data and the analytic model that utilize the data and the use of information systems to inform decision makers, and help them to get the advantage on their and histogram for giving the visual treat to the farmers. It provides the in-sights to the farmers about field. It is for creating the various visualization such a bar plot, bwplot, xyplot the growth of the crop for the better management of natural resources.

***Index Terms***—Data Analytics, Comma Separated Value, Regression, Prediction.

## I. INTRODUCTION

The basic idea of this project is the guide the farmers with the proper report about the crops which obtained by estimating the crop yields, cost of cultivation and the cost of production. The aim is to overcome all the drawbacks faced in all the existing applications and generate fast and accurate results. Though, we are living in a modernized world, agriculture plays a vital role.

Our project will be user-friendly as it is easily accessible by the users (i.e., farmers). As this project, will be the eye opener for the farmers and also for students who are taking courses about agriculture .In agriculture sector where farmers and

agribusinesses have to make the innumerable decisions every day and intricate complexities by involving the various factors influencing them. Our analytical report will help the farmers to make critical farming decisions. In this we are using various functions like `bwplot()` , `bar plot()` , `xypplot()` , which was supported by various packages like `lattice` for visualization , `dplyr` , `grid Extra` , `RcolorBrewer` . Separate analytical graph for each crops based on the states has been executed .Thus it also helps the government on providing the loans and insurance for the crops to the farmers for the safety, security and the well-being of the farmers.

## II. PROJECT DESCRIPTION

Our statistical analysis is the best solution for the farmers for improving their crop production and also to increase the crop yield by reducing the cost which is used for cultivation. One of the data mining algorithms has been used to make the statistics called K-means clustering algorithm .Nowadays, Data mining techniques plays vital role in the analysis of data. In this project ,we have been doing the analysis using R language , which a programming language and software environment for statistical analysis ,graphical representation and reporting and it is freely under the GNU General Public License, and pre-compiled binary versions are provided for various operating systems like Linux ,Windows and Mac. First download and install R and get the most useful package for machine learning in R. Load dataset and understand its structure using statistical summaries and data visualization .The process will be as same as we know defining the problem ,preparing data , evaluate algorithm , improve results , present results .In our project there are various modules, first module is about collecting the datasets and estimating the crop yield .second module is about analyzing cost of cultivation and cost of production. Final module is ofmaking state wise crop production analysis using K -means clustering and the final output has been obtained.

### III. LITERATURE SURVEY

#### *a) Analysis of agriculture data using data mining techniques*

There are the different forecasting methodologies developed and evaluated by the researchers all over the world in the field of agriculture. Some of such studies are: Researchers like Ramesh and Vishnu Vardhan are analyze the agriculture data for the years 1965-2009 in the district East Godavari of Andhra Pradesh, India. Rain fall data is clustered into the 4 clusters by adopting the K means clustering algorithm. Multiple linear regressions (MLR) is the method used to model their linear relationship between a dependent variable and one or more independent variables. The dependent variable is rainfall and the independent variables are year, area of sowing, production. Purpose of this work is to be find suitable data models that achieve the high accuracy and a high generality regarding of the yield prediction capabilities. Bangladesh offers the several varieties of rice which has different cropping season. For the prior study of climate (effect on temperature and rainfall) in Bangladesh and its effect on agricultural production of rice has been done. Then this study was being taken into the regression analysis with temperature and rainfall. Temperature puts an adverse consequence of the crop production. This pre-processing has been done for each kind of the rice variety. In clustering, the different pre-processed table has been analyzed to find the sharable group of region based on the similar weather attribute.

#### *b) A survey and analysis of various agricultural crops classification techniques*

Mining is the way of providing and extracting some meaning information from the data so that the data can be classified and grouped easily and quickly. These mining algorithms can be applied into the various fields including classification of agricultural crops production. In the fields of the Data Mining various efficient algorithms are implemented for the classification of agricultural crops production. Here in this paper a survey of all the

existing techniques as well as their advantages and issues are to be discussed. Hence by analyzing the various advantages and issues a new and efficient technique for their classification of their agricultural crops production is proposed in future such as classification using Fuzzy Conclusion Tree by the Optimizing the Feature Withdrawal using the PSO-SVM (Particle Swarm Optimization with a Support Vector Machine). But in the agriculture sector, predictive data mining is mainly used. There are two main techniques namely classification and the clustering. Data mining techniques have been put forward the guidelines for making recommendations for site specific crop management.

#### IV. EXISTING SYSTEM

- Enormous parameter does not exist.
- Based on Apriori algorithm, the computational cost is high, frequent item sets cannot be extended.
- It focuses on the analysis of the agriculture data and finding optimal parameters to maximize the crop production using data mining techniques like PAM, CLARA, DBSCAN and Multiple Linear Regression

#### V. SYSTEM ARCHITECTURE

Here, we are taking the datasets sample for our analysis .In that data sets we have the various plenty of information about the crops. According to this analysis we are going to estimate the crop yield and cultivation cost. In the given raw data sets we are going to perform the analytics for the given data. In that we are using an algorithm called multiple regression why the multiple regression means just doing a concept of prediction so it just need a predictor and response variable . Finally to visualize the data we are using a various library packages like xyplot and bwplot. In the graphical visualization the final outcomes which the highest crop production over the state. so the state wise yield for each crop has been analyzed and the final output has been displayed.

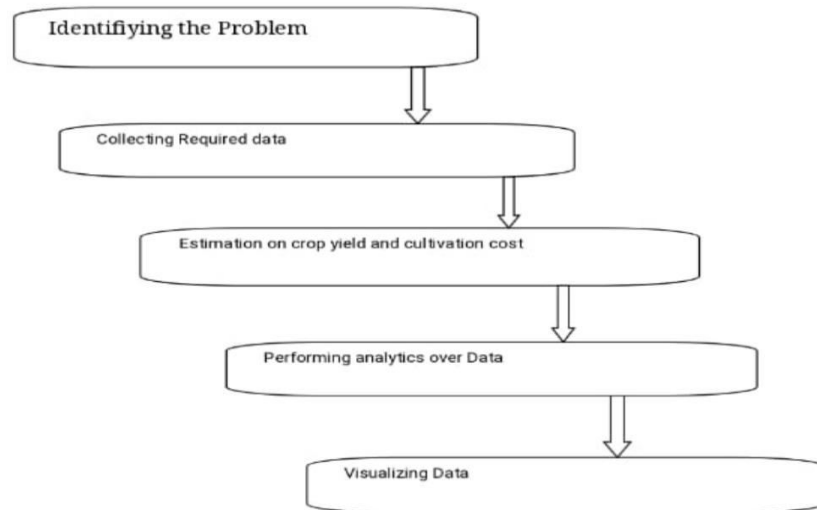


Fig. 1 System architecture for the crop estimation and cultivation

## VI. PACKAGES

### *Lattice and Lattice Extra:*

The lattice and lattice Extra add-on package is an implementation of Trellis graphics for R. It is a powerful and elegant the high-level data visualization system with an emphasis on multivariate data. It is designed to meet most typical graphic need with minimal tuning, but can also be easily extended to handle most nonstandard requirements.

### *Dplyr:*

Dplyr provides a flexible grammar of data manipulation. It's the next iteration of the plyr, focused on the tools for working with the data frames (hence the *d* in the name). Identifying the most important data manipulation verbs and make them easy to use from R.

*RcolorBrewer:*

Creating the nice looking color palettes especially for thematic maps.

RColorBrewer {RColorBrewer}

## VII. RESULT

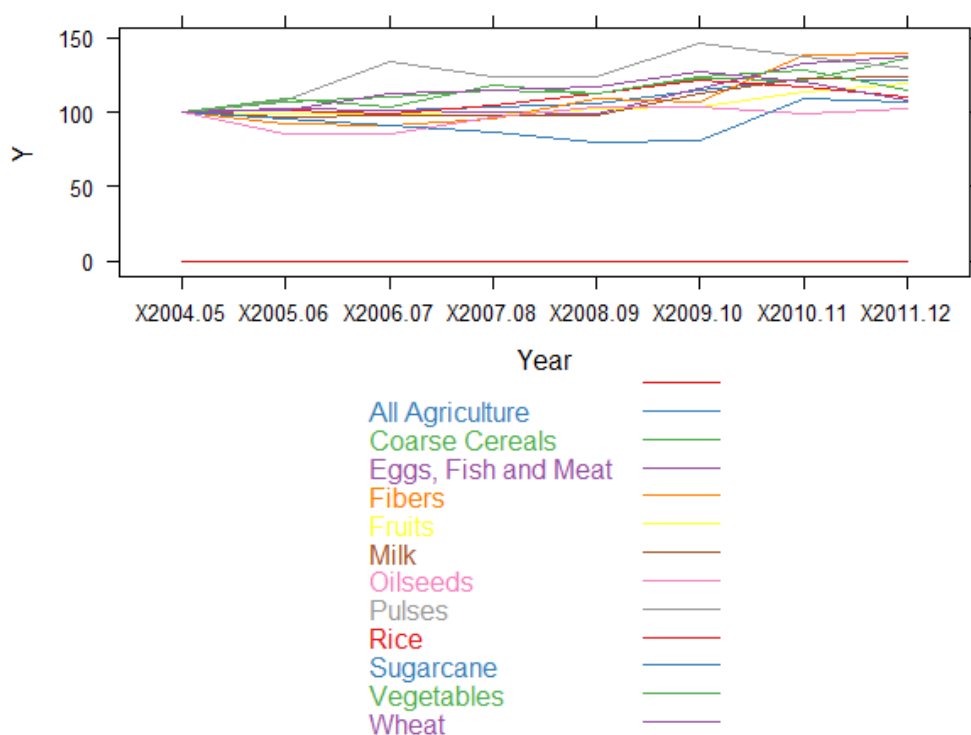


Fig .2 Result analysis for crop estimation

Dataset is taken from agricultural open source site. This project analysis the agricultural production results from the statistical point of view. Computing the crop production for the farmers’ growth. Creating the various visualization such as bar plots, BW plots, XY plots, Histogram and Perform for giving the visual treat to the farmers’. It is insight to the farmers’ for the better management resources.

## VIII. CONCLUSION

The purpose of this project is to analyze the relationships between the yield and cost of various crops based on the states. Various data mining techniques are to be implemented on the input data to assess the best performance yielding method. Clustering methods are compared using quality metrics. Several crops have been required for yield estimation and it is collected using the data availability as the key measure. Thus, the crop data over 8 years has been collected and analyzed using one of the data mining techniques. Finally, from the collected data crop estimation has been visualized to provide insights to farmers about the productivity of crops on different states.

## ACKNOWLEDGMENT

Our sincere thanks to the staff who helps us for the paper about the project.

## REFERENCES

- [1] Surabhi Chouhan, Anju Singh,, A survey and analysis of various agricultural crops classification techniques, *journal of big data published in 2016*.
- [2] Keerthana B, Nivetha P, Boomika M, Literature Review on IOT Based Smart Security and Monitoring Devices for Agriculture, International Journal on Future Revolution in Computer Science & Communication Engineering. February 18 Volume 4 Issue 2.
- [3] Jhama Majumdar, Sneha Naraseeyappa , Analysis of agriculture data using data mining techniques, *Journal of Big Data 2017*.
- [4] G. Nasrin Fathima, R.Geetha , AGRICULTURE CROP PATTERN USING DATA MINING TECHNIQUES International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 5, May 2014.
- [5] Ajay kakra, Trends in Agricultural, International Journal Big data 2016.
- [6] Ramalatha Marimuthu, Kanagaraj. S, Design and Development of a Persuasive Technology Method to encourage Smart Farming 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)21 - 23 Dec 2017, Dhaka, Bangladesh
- [7] Apeksha Thorat, Sangeeta Kumari An IoT Based Smart Solution for Leaf Disease Detection 2017 International Conference on Big Data, IoT and Data Science (BIG DATA) Vishwakarma Institute of Technology, Pune, Dec 20-22, 2017
- [8] M. R. Bendre, R. C. Thool Big Data in Precision Agriculture: Weather Forecasting for Future Farming 1st International Conference on Next Generation Computing Technologies (NGCT-2015) Dehradun, India, 4-5 September 2015
- [9] JHARANA MANJUDHAR, Analyses of agriculture in smart farming in International Journal in Big data <https://www.iaescore.com/journals/index.php/IJEECS/article/view/13056/9927>
- [10] RAMYA M.G, CHETAN BALAJI , Environment change prediction in Smart Agriculture using big data analytics. <https://ieeexplore.ieee.org/document/8342819/keywords>

**AUTHOR(S) DETAILS:**



P. Sakthivel, AP/CSE, Dhirajlal Gandhi College of Technology, Salem, Tamilnadu



S. S. Aravinth, AP/CSE, Dhirajlal Gandhi College of Technology, Salem, Tamilnadu



R. Radhapriya, UG Scholar, CSE, Dhirajlal Gandhi College of Technology, Salem, Tamilnadu



M. Sangeetha, UG Scholar, CSE, Dhirajlal Gandhi College of Technology, Salem, Tamilnadu



K. Pavithra, UG Scholar, CSE, Dhirajlal Gandhi College of Technology, Salem, Tamilnadu