

DATA ANALYSIS OF CROP YIELD PREDICTION USING K-MEANS CLUSTERING ALGORITHM

¹Dr.R.Umarani , ²P.Tamilarasi

¹Associate professor, ²Research Scholar

¹Department of Computer Science,

^{s1}. Sri Sarada College for Women, Salem-16, India.

Abstract: Agriculture is the backbone of India. But now Indian agriculture is in underprivileged condition. Because the crop yield depends upon the monsoon rain fall and various other factors and those factors do not favor the production of the crops to the maximum. Crop yield prediction is an extremely interested research part in agriculture. Yield prediction is very tricky to be solved based on the presented data. Yield prediction helps the farmers to plan the requirements in advance. Their various deeds are based on only crop yields. Hence yield prediction is very useful to the farmers. Using data analytics techniques we can crack this problem efficiently. Data analytics plays an important role to examine data sets in various fields such as agriculture, medical, crime, commercial industries etc. In this paper we proposed a brief analysis of crop yield prediction using k-means clustering techniques based on different parameters like years, areas and production of various crops in India. The k-means cluster algorithm is used to locate group in the data, with the figure of group represented by the variable K. This algorithm is planned to give all the data points to one of K group based on the characteristic. This algorithm was developed and implemented using R.

Index Terms: Crop Yield, Data analytics, K-means clustering, Prediction.

I. INTRODUCTION

Farmers are not receiving predictable crop yield due to number of reasons in India. The yield mainly depends on climate. Rainwater also manipulates the yield farming. In this circumstance, the farmers essentially require a sensible guidance to envisage the future crop productivity and an examination is to be made in order to assist the farmers to exploit the crop manufacture in their crops. Yield calculation is an essential agricultural dilemma. Yield prophecy was executed by in view of farmer's earlier experience on a scrupulous crop. Data analytics is mostly applied to agricultural tribulations. Data analytics is used to analyze big data sets and launch valuable classifications in the dataset. The overall target of the data analytics procedure is to extort the information from a records set and transform to explicable structure for further use. In this paper the aspire is to build a user gracious interface for farmers, which gives the analysis of crop production pedestal on existing data. Data analytics K-means cluster techniques were used to predict the crop yield for make the most of the crop productivity.

II. LITERATURE SURVEY

The data of quantity has been rising and dataset examined develop into more aggressive. The combination of analytics techniques and decision optimizations are called predictive analytics. The summary of evolutionary algorithms is executed by descriptive survey and analysis [1]. The analysis of the agriculture data and pronouncement optimal variables to exploit the crop production with mining techniques similar to PAM, CLARA, DBSCAN and Multiple Linear Regression. [2].

This proposed system focused to gain approaching into the Crop Yield forecast with Big Data Analysis and recognize the related socio-economic dispute. In this development the analysis of vast data would be accepted using K-means clustering methods to analyze the best appropriate way of agriculture methods in that exact region and forecast of yield would be established by Apriori algorithms and this valuable data would be over and again specified to farmers for the enhanced results of crop yield [3].

This research proposes and applies a system to predict crop yield from earlier data. This is accomplished by concern association rule mining on agriculture data [4]. At first the statistical model Multiple Linear Regression technique is used on previous data. The outcomes so achieved were tested and examined using the Data Mining technique that is to say Density-based clustering technique. In this method the outcome of two methods were matched up to according to the exact region [5]. K-means algorithm is used for prediction. Because forecast rainfall data is based on year by year [6].

This paper is focused to study various clustering techniques that are integrated on the seed data sets to develop the clustering approach support on different factors [7]. The aim of the paper is to calculate the crop yield using density based clustering techniques. This clustering techniques is establish proper for the approximate prediction.[8]. In this previous paper discuss about data mining technique centric support clustering was urbanized and to conclude the vast production of KAHIRIF and RABI [9]. Data mining functionalities that are used to identify the variety of patterns to be originate in data mining task [10]

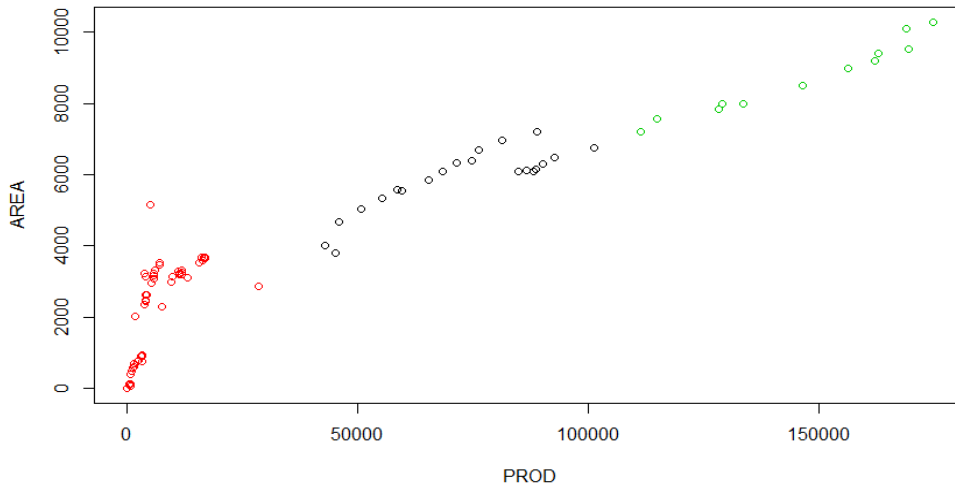


Figure2: Area wise production using k-means cluster plot

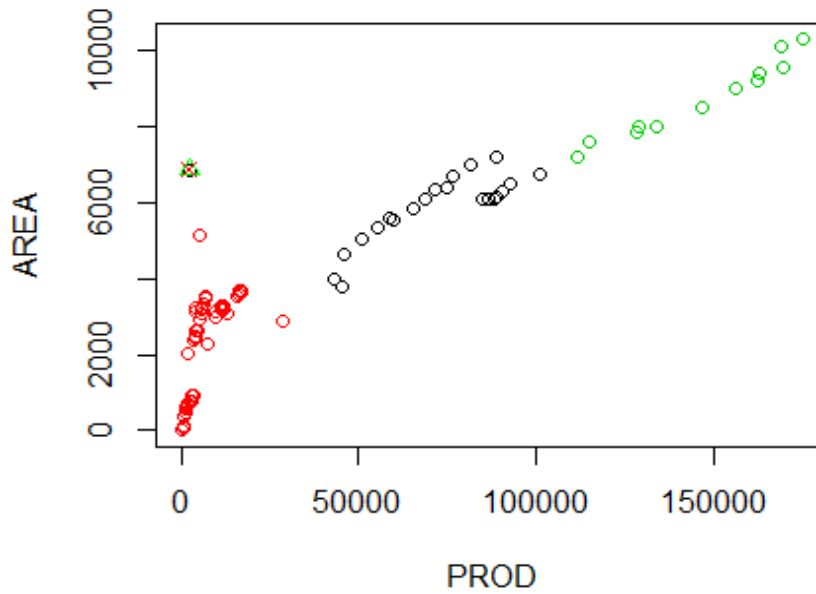


Figure 3: Result of center cluster

S.No	Area	Production
1	2298.077	6864.769
2	8719.250	146597.417
3	5881.762	72297.381

Table2: Result of centre cluster values

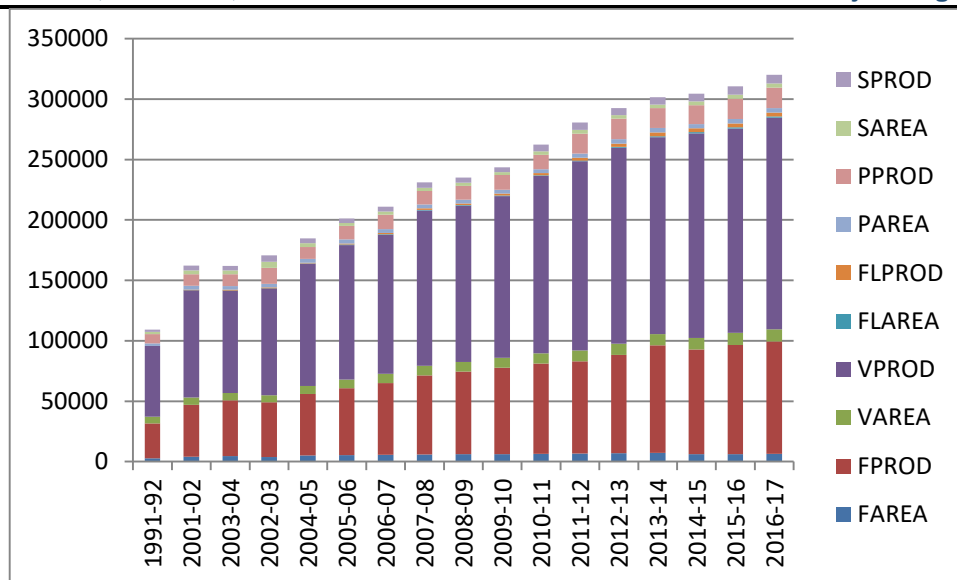


Figure 4: Year Wise Crop Production

The dataset of crops as specified in Table1. Figure1 shows the cluster of crop production. It specified high level cluster in green. 2ndCluster is specifying average stage of crop production in black.1st Cluster is specifying poor stage of crop making in red. Figure 3 and Table2 shows the cluster centre point and cluster centre values. Figure 4 shows the production of crops based on the year and area. The production of vegetable (VPROD) is highest than other crops in every year.

V.CONCLUSION

In this paper k-means clustering technique was used to predict and determine the agricultural crop yield. This algorithm predicts the production of crop for a specific period in India. This work helps the farmers to get awareness on identifying less productive crop yield in large area. Though there are various clustering techniques existing in data analytics, K-means clustering technique is proper and found to give fairly accurate prediction. In the consequential work analysis, prediction of crop production can be developed by different data analytics techniques like Support Vector Machine (SVM), K-Nearest Neighbor.

REFERENCE

- [1]S .Banumathi, A.Aloysius,” Big data prediction using evolutionary techniques: a survey”, jetir September 2016.
- [2] [1] Majumdar, J., Naraseeyappa, S., & Antalkali, S. (2017). *Analysis of agriculture data using data mining techniques: application of big data. Journal of Big Data.*
- [3] Ms. Ruchita Thombare, Ms. Shreya Bhosale, Mr. Prasanna Dhemy, Ms.Anagha Chaudhari, “Crop Yield Prediction Using Big Data Analytics”, IJCMS November 2017
- [4] E.Manjula, S.Djodiltachoumy Pachaiyappas, “A Model for Prediction of Crop Yield”, International Journal of Computational Intelligence and Informatics, Vol. 6: No. 4, March 2017
- [5] D Ramesh , B Vishnu Vardhan, “ Analysis of crop yield prediction using data mining techniques “ijret Jan-2015
- [6] Awanit Kumar , ShivKumar, ” Prediction of Production of Crops using K-mean & Fuzzy Logic”, IJCSMC, Vol. 4, Issue. 8, August 2015, pg.44 – 56
- [7] Madhavi Gudavalli , Vidyasree P , S Viswanadha Raju “Clustering Analysis for Appropriate Crop Prediction using Hierarchical, Fuzzy C-Means, K-Means and Model based Techniques”, IJAERD-2017
- [8] B Vishnu Vardhan and D. Ramesh,” Density Based Clustering Technique on Crop Yield Prediction”, International Journal of Electronics and Electrical Engineering Vol. 2, No. 1, March, 2014
- [9] Raj Kumar Tripathi, Nishtha Kesswani,” Clustering the Indian States on the Basis of Agriculture Produce of KHARIF and RABI Crops”, International Journal of Electronics Communication and Computer Technology (IJECCT) Volume 2 Issue 2 (March 2012)
- [10]Utkarsha P. Narkhede,2K. P. Adhiya,” A Study of Clustering Techniques for Crop Prediction - A Survey, AIJRSTEM - 2014