[Volume 3 issue 11 November 2015]
PageNo.745- 750 ISSN : 2320-7167

INTERNATIONAL JOURNAL OF MATHEMATICS AND COMPUTER RESEARCH

An Analysis Of Crop Return Probability Consuming Using Data Mining Techniques

P.Sharmila Devi, M.C.A., M.Phil.,[1] Dr. R.Periyasamy M.Sc., (Phy), PGDCA, M.Sc., M.C.A., M.Phil., Ph.D.,[2]

[1]Research scholar, Dept. of CS, Nehru Memorial College (autonomous) Puthanampatti, Tiruchirappalli.
sharmimca.112@gmail.com

[2]Associate professor, Dept. of CS, Nehru Memorial College (autonomous), Puthanampatti, Tiruchirappalli.
rpsamy64@yahoo.com

Abstract : Agrarian sector in India is facing rigorous issue to augment the crop productivity. More than 60 percent of the crop still depends on monsoon rainfall. Recent developments in Data Innovation for farming field has become an interesting research zone to anticipate the crop yield. The issue of yield expectation is a major issue that remains to be solved based on accessible data. Data Mining strategies are the better choices for this purpose. Diverse Data Mining strategies are utilized also, evaluated in farming for estimating the future year's crop production. This paper presents a brief examination of crop yield expectation utilizing Various Straight Relapse (MLR) system also, Thickness based grouping system for the selected zone i.e. East Godavari zone of Andhra Pradesh in India.

Keywords: Agrarian Sector, Crop Production, Data Mining, Thickness based clustering, Data Technology, Various Straight Regression, Yield Prediction.

1. Introduction

Farming is the backbone of Indian Economy. In India, majority of the agriculturists are not getting the expected crop yield due to several reasons. The farming yield is primarily depends on atmosphere conditions. Precipitation conditions moreover influences the rice cultivation. In this context, the agriculturists necessarily requires a timely advice to anticipate the future crop proficiency also, an examination is to be made in request to help the agriculturists to augment the crop generation in their crops.

Yield expectation is an important farming problem. Every farmer is interested in knowing, how much yield he is about expect. In the past, yield expectation was performed by considering farmer's previous experience on a specific crop. The volume of data is enormous in Indian agriculture. The data when become data is highly helpful for many purposes.

Data Mining is generally connected to farming problems. Data Mining is utilized to analyze expansive data sets also, establish helpful characterizations also, patters in the data sets. The overall goal of the Data Mining process is to extract the data from a data set also, transform it into understandable structure for further use.

In this paper the main aim is to create a user friendly interface for farmers, which gives the examination of rice generation based on accessible data. Diverse Data mining strategies were utilized to anticipate the crop yield for maximizing the crop productivity.

2. Literature Survey

From the research article, the analyst express that expansive amount of data which is collected also, stored for analysis. Making appropriate use of these data often leads to considerable gains in proficiency also, therefore economic advantages.
There are several applications of Data Mining strategies in the field of agriculture. The specialists implemented K-Means algorithm to forecast the pollution in the atmosphere, the K nearest Neighbour is connected for simulating daily precipitations also, other atmosphere variables also, diverse possible changes of the atmosphere scenarios are investigated utilizing Support Vector Machines.

Soil profile descriptions were proposed by the analyst for characterizing soils in combination with GPS based technologies. They were connected K-Means approach for the soil classification. In a similar approach, crop characterizations utilizing hyper spectral data was conveyed out by adopting one of the data mining approach i.e. Support Vector Machines. One of the analyst utilized an intensified fuzzy group examination for characterizing plants, soil also, residue regions of interest from GPS based colour images.

In the farming science, grouping strategies are found in grading apples before marketing. Weeds were detected on precision agriculture. The specialists worked on precipitation variability examination also, its sway on crop productivity. The effect of observed seasonal climatic conditions such as precipitation also, temperature variability on crop yield expectation was considered through an empirical crop model. Furthermore, there are two approaches to investigate the sway of atmosphere change on crop generation which include the crop suitability approach also, the generation function approach.

Specialists were found that the yields of winter wheat are reduced when temperatures rise, due to the consequent reduction of the growth phases of the plant also, moreover concluded that the complexity of a model was based on the level of detailed examination or it was less detailed with just estimations of moisture content.

3. Overview of Data

The data utilized for this paper are acquired for the years from 1955 to 2009 for East Godavari zone of Andhra Pradesh in India. The preliminary data accumulation is conveyed out for all the areas of Andhra Pradesh in India. Each zone in this accumulation is identified by the respective longitude also, latitude of the region. The assessment is considered for just East Godavari zone of Andhra Pradesh in India.

The data are taken in eight input variables. The variables are 'Year', 'Rainfall', 'Zone of Sowing', 'Yield', 'Fertilizers' (Nitrogen, Phosphorous also, Potassium) also, 'Production'. The trait 'Year' determines the year in which the data are accessible in Hectares. 'Rainfall' trait determines the average precipitation in the determined year in Centimeters. 'Zone of Sowing' trait determines the total zone sowed in the determined year for that zone in Hectares. 'Yield' determines in Kilogram per hectare. 'Production' trait determines the generation of crop in the determined year in Metric Tons. 'Fertilizers' specify in Tons in the determined year.

4. Methodology

In this paper the factual system namely various straight relapse system also, Data mining system namely Density-based grouping system were taken up for the estimation of crop yield analysis.

4.1 Various Straight Regression

A Relapse model that involves more than one predictor variable is called Various Relapse Model. Various Straight Relapse (MLR) is the method, utilized to model the Straight relationship between a subordinate variable also, one or more insubordinate variables. The subordinate variable is sometimes termed as predicting also, insubordinate variables are called predictors.

Various Straight Relapse (MLR) system is based on minimum squares also, probably the most generally utilized system in climatology for developing models to reconstruct atmosphere variables from tree ring services. This crop yield expectation model is presented with the use of Various Straight Relapse (MLR) system where the predictant is the Generation also, there are seven indicators namely Year, Rainfall, Zone of Sowing, Yield also, Manures (Nitrogen, Phosphorous also, Potassium).

4.2 Density-based Grouping Technique

The primary thought of Density-based grouping strategies is that, for each point of a cluster, the neighborhood of a given unit distance contains at minimum a minimum number of points. In other words the Thickness in the neighborhood should reach some threshold. However, this thought is based on the assumption that the bunches are in the spherical or regular shapes.

These techniques group the objects concurring to specific Thickness objective functions. Thickness is usually defined as the number of objects in a specific neighborhood of data objects. In these approaches, a given group continues to grow as long as the number of objects in the neighborhood which exceeds some parameter. This is considered to be diverse from the thought in partitioning algorithms that use iterative relocation of focuses that give a certain number of clusters.

5. Results Also, Discussion

In this paper an effort is made in request to know the zone specific crop yield examination also, it is processed by implementing both Various Straight Relapse system also, Density-based grouping technique. These models were experimented in respect of all the areas of Andhra Pradesh, but the process of assessment is conveyed out with just East Godavari zone of Andhra Pradesh in India.

The Accurate esteem along with the relating evaluated esteem utilizing Various Straight Relapse system for 40 years interim of test data about East Godavari Zone is appeared in the Table-1.

The evaluated results utilizing Various Straight Relapse system which are extending between -14% also, +13% for 40 years interval.

Table-1: Accurate generation also, evaluated values utilizing Various Straight Relapse technique.

Observation Year	Production (Accurate)	40 Years Interval	
		Production (Estimation)	Percentage of Difference
2000	683423	592461	13
2001	579850	566050	2
2002	551115	579433	-5
2003	762453	722638	5
2004	743614	742752	0
2005	348727	399062	-14
2006	547716	551541	-1
2007	715472	691069	3
2008	716609	697227	3
2009	616567	633494	-3

The estimation of the crop yield expectation utilizing Density-based grouping system for 6-groups guess of test data about East Godavari Zone is appeared in the Table-2. The evaluated results utilizing Density-based grouping system which are extending between -13% also, +8% for 6-groups approximation.

Table-2: Accurate generation also, evaluated values utilizing Density-based grouping technique

Observation Year	Production (Accurate)	6 Clusters	
		Production (Estimation)	Percentage of Difference
2000	683423	666011	3
2001	579850	651103	-12
2002	551115	566972	-3
2003	762453	703914	8
2004	743614	737897	1
2005	348727	392770	-13
2006	547716	534709	2
2007	715472	791589	-11
2008	716609	676321	6
2009	616567	695574	-13

The examination between Accurate generation along with the relating evaluated esteem utilizing Various Straight Relapse system for 40 years interim also, Density- based grouping system for the 6-groups guess about East Godavari Zone is appeared in the following Table-3 also, Figure-1.

Table-3: Examination between Accurate generation also, evaluated values utilizing Various Straight Relapse system also, Density-based grouping technique

Observation Year	**Production (Accurate)**	**Generation (Estimation)**	
		Multiple Linear Regression Technique	**Density-based clustering technique**
2000	683423	592461	666011
2001	579850	566050	651103
2002	551115	579433	566972
2003	762453	722638	703914
2004	743614	742752	737897

2005	348727	399062	392770
2006	547716	551541	534709
2007	715472	691069	791589
2008	716609	697227	676321
2009	616567	633494	695574

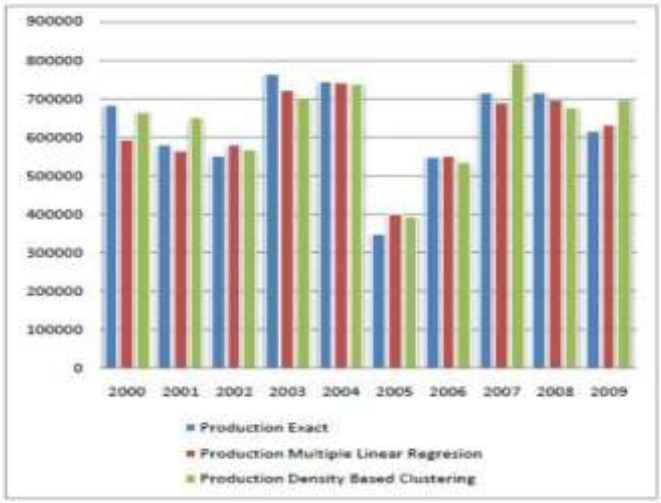

Fig-1: Examination between Various Straight Relapse system also, Density-based grouping technique

5. Conclusion

Initially the factual model Various Straight Relapse system is connected on existing data. The results so acquired were verified also, investigated utilizing the Data Mining system namely Density-based grouping technique. In this procedure the results of two techniques were compared concurring to the specific zone i.e. East Godavari zone of Andhra Pradesh in India. Similar process was adopted for all the areas of Andhra Pradesh to improve also, authenticate the validity of yield expectation which are helpful for the agriculturists of Andhra Pradesh for the expectation of a specific crop. In the subsequent work a examination of the crop yield expectation can be made with the entire set of existing accessible data also, will be dedicated to suitable approaches for improving the proficiency of the proposed technique.

References

[1] Leigang Shi; Coll. of Agronomy & Biotechnol., China Agric. Univ., Beijing, China; Zaijie Yang; Fu Chen, "The carbon sequestration capacity in crop production system in North China Plain", Published in: Electrical and Control Engineering (ICECE), 2011 International Conference on Date of Conference: 16-18 Sept. 2011 Page(s): 4697 – 4700.

[2] Jianbing Zhang; State Key Lab. of Remote Sensing Inf., Chinese Acad. of Sci., Beijing, China; Jinping Li; YePing Zhu, "Integrating crop simulation models with WebGIS for remote crop production management", Published in: Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International (Volume:4) Date of Conference: 20-24 Sept. 2004 Page(s): 2598 - 2600 vol.4.

[3] Paswan, R.P.; Dept. of Comput. Sci., Assam Univ., Silchar, India; Begum, S.A., "ANN for prediction of Area and Production of Maize crop for Upper Brahmaputra Valley Zone of Assam", Published in: Advance Computing Conference (IACC), 2014 IEEE International Date of Conference: 21-22 Feb. 2014 Page(s): 1286 – 1295.

[4] Chaochao Gao; Dept. of Environ. Sci., Zhejiang Univ., Hangzhou, China, "The impact of climate change on China's crop production: A CMIP5 ensemble assessment", Published in: Agro-Geoinformatics (Agro-Geoinformatics), 2012 First International Conference on Date of Conference 2-4 Aug. 2012 Page(s): 1 – 5.

[5] Guo-Qiang Yuan; Fundamental Dept., Hebei Finance Univ., Baoding, China; Yu-hui Liu; Can Cao, "Fuzzy crop production planning expected value model with credibility constraints", Published in: Fuzzy Systems and Knowledge Discovery (FSKD), 2011 Eighth International Conference on (Volume:1) Date of Conference: 26-28 July 2011 Page(s): 30 – 634.

[6] Zagar, M.; Polytech. of Zagreb, Zagreb, Croatia; Poljak, D.M., "Have we been monetized and become commodity without our consent - Privacy in the time of Big Data technology", Published in: Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2015 38th International Convention on Date of Conference: 25-29 May 2015 Page(s): 1344 – 1349.

[7] Shangyi Chen; Univ. of Chinese Acad. of Sci., Beijing, China; Wei Li; Min Li; Xiaofei Zhang, "Latest Progress and Infrastructure Innovations of Big Data Technology", Published in: Cloud Computing and Big Data (CCBD), 2014 International Conference on Date of Conference: 12-14 Nov. 2014 Page(s): 8 – 15.

[8] Panahiazar, M.; Center for Sci. & Healthcare Delivery, Mayo Clinic, Rochester, MN, USA; Taslimitehrani, V.; Jadhav, A.; Pathak, J., "Empowering personalized medicine with big data and semantic web technology: Promises, challenges, and use cases", Published in: Big Data (Big Data), 2014 IEEE International Conference on Date of Conference: 27-30 Oct. 2014 Page(s): 790 – 795.

[9] Chanmin Park; Kim, Lee & Park Law Offices, Beverly Hills, CA, USA; Taehyung Wang, "Big Data and NSA Surveillance -- Survey of Technology and Legal Issues", Published in: Multimedia (ISM), 2013 IEEE International Symposium on Date of Conference: 9-11 Dec. 2013 Page(s): 516 – 517.