International Journal of Information & Computation Technology.
ISSN 0974-2239 Volume 6, Number 1 (2016), pp. 25-33

http://www. irphouse.com

Crop Yield Prediction with Aid of Optimal Neural Network in Spatial Data Mining: New Approaches

Aakunuri Manjula[1] and Dr. G.Narsimha[2]

[1]Research Scholar, CSE, JNTUH, Telangana State, India.
E-mail: manjula3030@gmail.com

[2]Associate Professor, CSE, JNTUH, Telangana State, India.
E-mail: narsimha06@gmail.com

Abstract

Data Mining is the process of extracting useful information from large datasets. Data mining techniques till now used in business and corporate sectors may be used in agriculture for data characterization, discrimination and predictive and forecasting purposes. Data mining in agriculture is a novel research field. Recently Knowledge Management in agriculture facilitating extraction, storage, retrieval, transformation, discrimination and utilization of knowledge in agriculture. Agriculture data are highly expanded in provision of nature, interdependencies and resources. The agriculture yield is primarily depends on weather conditions, diseases and pests, planning of harvest operation, geographical and biological factors and the likes. As far as data mining techniques is concern in the most of cases predictive data mining approaches is used. Predictive data mining is used to predict the future crop, weather forecasting, pesticides and fertilizers to be used, revenue to be generated and so on. Crop yield prediction has been a topic of interest for producers, consultants and agricultural related organizations. As defined by the Food and Agriculture of the United Nations, crop forecasting is the art of predicting crop yields and production before harvest takes place, typically a couple of months in advance.

Keywords: Spatial data mining, Crop yield, artificial neural network, multilinear principal component analysis, prediction.

I. INTRODUCTION

The goal of the data mining process is to extract knowledge from an existing data set and transform it into a human understandable formation for advance use [1]. It is the process of analyzing data from different perspectives and summarizing it into useful information. There is no restriction to the type of data that can be analyzed by data mining. As far as data mining technique is concern; in the most of cases predictive data mining approach is used. Predictive data mining technique is used to predict future crop, weather forecasting, pesticides an d fertilizers to be used, revenue to be generated and so on [2]. Data mining is in great demand in other applicative databases, including spatial databases, temporal databases, object-oriented databases, multimedia databases, etc. Spatial data mining can be used to predict crop yield prediction besides other real world applications. Spatial data mining is the process of extracting interesting knowledge from spatial databases. This knowledge can be used for understanding spatial and non- spatial data and their relationships. Knowledge discovered from spatial data can be of various forms, like characteristic and discriminant rules, extraction and description of prominent structures or clusters, spatial associations, and others [3]. The challenge has been to extract knowledge from this raw data; this has lead to new methods and techniques such as data mining that can bridge the knowledge of the data to the crop yield estimation [4].

Crop yield prediction has been a topic of interest for producers, consultants, and agricultural related organizations. As defined by the Food and Agriculture of the United Nations, crop forecasting is the art of predicting crop yields and production before the harvest actually takes place, typically a couple of months in advance [5]. Crop yield prediction is a significant component of national food security assessment and food policy making. Crop growth and yield data are critical for regulating agricultural cultivation system, and agricultural operation and management [6]. The crop production could be affected directly and indirectly under climate change [7]. Statistical models and crop models are two main tools for studying effects of climate change on crop yields [8]. Majority of research works in agriculture focus on biological mechanisms to identify crop growth and improve its yield. The outcome of crop yield primarily depends on parameters such as variety of crop, seed type and environmental parameters such as sunlight, soil, water, rainfall and humidity [9]. Crop yield estimation is necessary, particularly in countries like India, which depends on agriculture as their main source of economy. Crop yield estimation helps decisions makers to take decision in respect to the excess or deficit production conditions and allow timely import and export decisions [10]. Crop identification and prediction of yield are the main concern of remote sensing application in agriculture [11].

II. LITERATURE REVIEW

Several techniques were proposed by various authors for crop yield prediction and a few of them are explained below:

E.I. Papageorgiou et al [12] have investigated the process of yield prediction in cotton crop production using the soft computing technique of fuzzy cognitive maps. Fuzzy cognitive map (FCM) was a fusion of fuzzy logic and cognitive map theories, and was used for modeling and representing experts' knowledge. It was capable of dealing with situations including uncertain descriptions using similar procedure such as human reasoning does. It was a challenging approach for decision making especially in complex processing environments. Xiaoqin Dai et al[13] have proposed simulation for response of crop yield to soil moisture and salinity with artificial neural network. In saline fields, irrigation management often requires understanding crop responses to soil moisture and salt content. Luke Bornn and James V. Zidek [14] have describes how spatial dependence could be incorporated into statistical models for crop yield along with the dangers of ignoring it. In particular, approaches that ignore this dependence suffer in their ability to capture the underlying phenomena. By judiciously selecting biophysically based explanatory variables and using spatially-determined prior probability distributions, a Bayesian model for crop yield was created that not only allows for increased modeling flexibility but also for improved prediction over existing least-squares methods.

The impact of climate change on cotton yields in seven main arable crop sites in Greece was investigated by Dimitrios Voloudakis et al [15]. The FAO AquaCrop (v.4) water driven model was used as a crop growth simulation tool under eight climatic models based on IPPC's A1B emission scenario. Predicting annual crop yields is of interest for many agricultural applications. David Gouache et al [16] have presented a prediction scheme at the departmental level, circa 100 km by 100 km, of winter wheat yields in France, applied for 23 departments, using official yield statistics from 1986 to 2010. Each model was a linear combination of 5–7 variables, selected from an initial pool of over 250 candidates. Variable selection was carried out with forward stepwise regression methods. The variable selection process was cross-validated, so as to select robust variables. Model prediction performance was also evaluated by cross-validation.\

III. PROBLEM DEFINITION

Early prediction of crop yield is important for planning and taking various policy decisions. Many countries use the conventional technique of data collection for crop monitoring and yield prediction based on ground based visits and reports. These methods are subjective, very costly and time consuming. The common problem in existing crop yield prediction methods are given below,

- The most important problem of existing crop yield prediction method is accuracy and time consuming problem.
- In existing time series crop yield prediction method does not react to variations that occur for cycles and seasonal effects.
- Needs extensive information to develop and test the model and also available information in agriculture is sparse and incomplete in existing simulation model.
- Limited studies have been made in crop yield prediction using existing decision tree technique.
- Prediction error value also important problem in crop yield prediction or estimation methods.
- These are the main drawbacks of various existing works, which motivate us to do this research on crop yield prediction.

IV. METHOD OF ANALYSIS

Crop yield prediction can help agricultural departments to have strategies for improving agriculture. Crop production depends on climatic, geographical, biological, political and economic factors. Because of these factors there are some risks, which can be quantified when applied appropriate mathematical or statistical methodologies. Actually accurate information about the nature of historical yield of crop is important modeling input, which are helpful to farmers & Government organization for decision making process in establishing proper policies. In this paper we have intend to propose a method for crop yield prediction using classifier. The proposed crop yield prediction consists of three phases namely, preprocessing, feature reduction and prediction. Here the proposed method use input data as real world data. Real world data is often incomplete, inconsistent, and/or lacking in certain behaviors or trends, and is likely to contain many errors. Data pre-processing is a proven method of resolving such issues. A good data preprocessing helps to create better model and will consume less time. Next phase of the proposed method is feature reduction; here we use multilinear principal component analysis (MPCA) for feature reduction phase. Finally the proposed method is use to predict the crop yield by means of classifier. The prediction is done based on the Optimal Neural Network classifier (ONN). The performance of the proposed method is evaluated by prediction accuracy and error value.

V. RESULTS AND DISCUSSION

Agriculture in Telangana state reflects many crops being cultivated. The cultivated land is distributed across the state. Food and Non-food crops are cultivated. The area in which particular crop or crops are cultivated is shown in Figure 1. It is evident that the crops such as Cotton, Maize and so on are cultivated in different areas of Telangana state. However, in this Telangana state area is considered for experiments.

The data is collected from officials of agriculture department of Telangana, India. The data is collected for years from 2011 to 2015

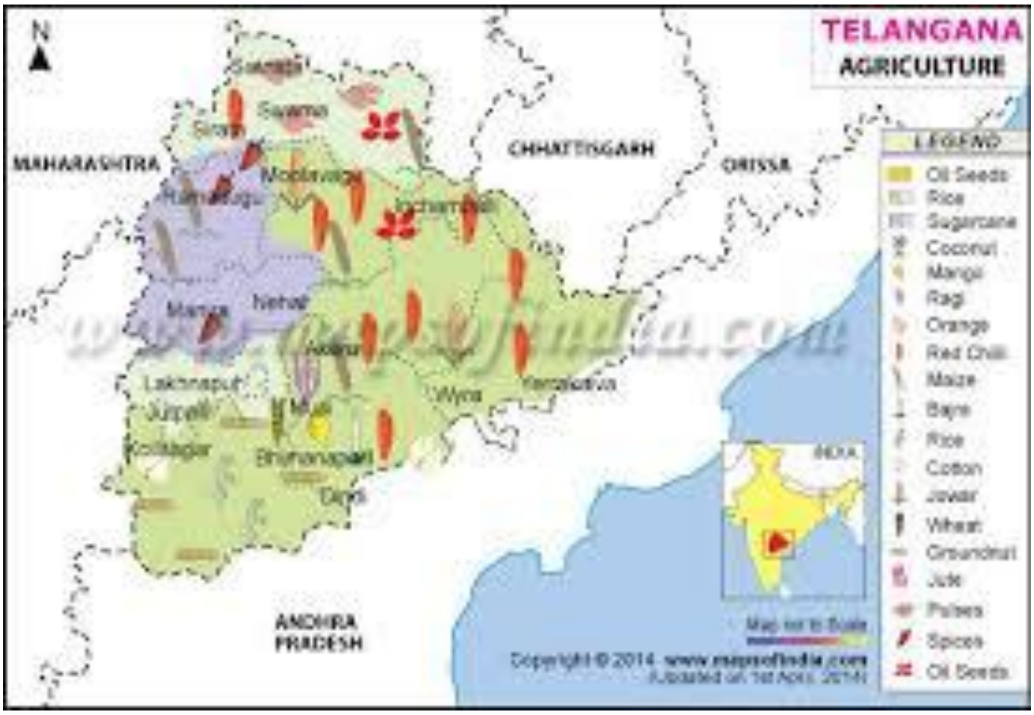
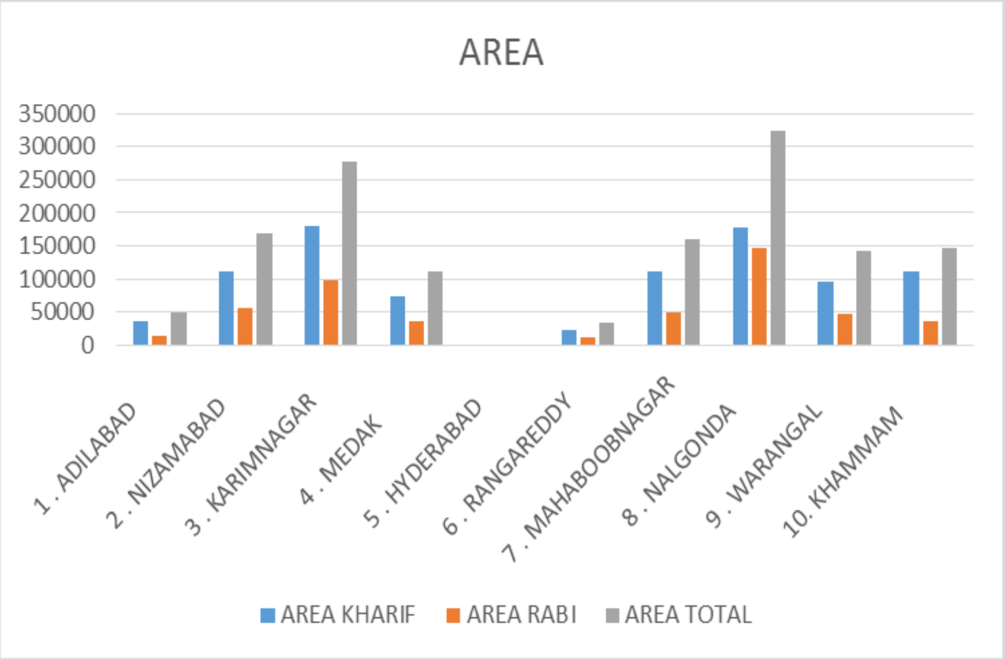


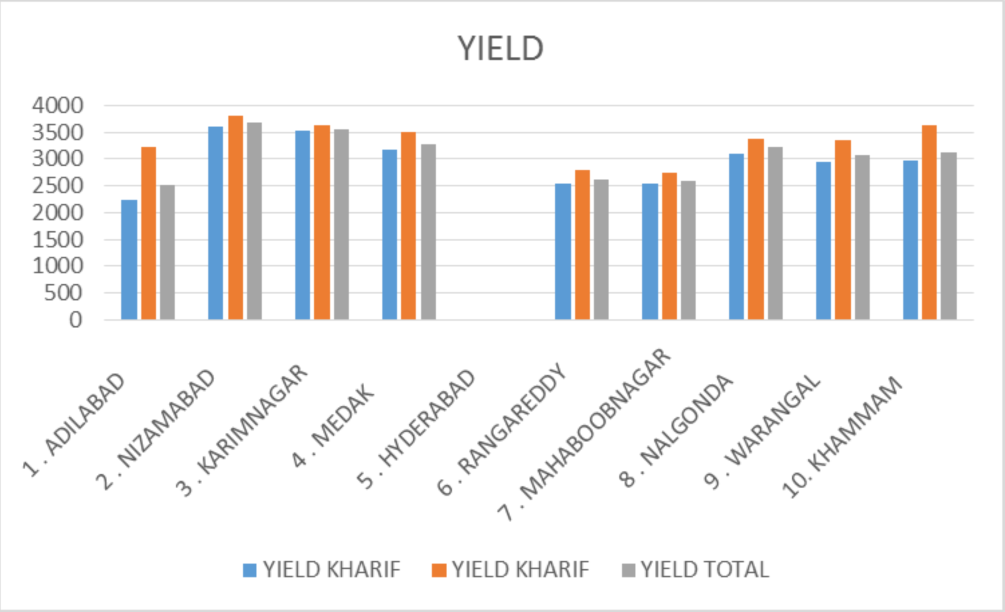
Figure 1: Agriculture crops dynamics in Telangana

AREA, YIELD AND PRODUCTION REPORT FOR THE YEAR 2015-16
CROP: RICE (Area in Hectares) (Yield in Kgs/Hect.)(Production in Tonnes)

Table 1: Area, Yield and Production Report for Rice Crop

DISTRICT	AREA			YIELD			PRODUCTION		
	KHARIF	RABI	TOTAL	KHARIF	RABI	TOTAL	KHARIF	RABI	TOTAL
1. ADILABAD	42575	14929	57504	2550	2882	2636	108571	43025	151596
2. NIZAMABAD	80451	17578	98029	2595	3073	2681	208770	54017	262787
3. KARIMNAGAR	116345	54053	170398	3277	3195	3251	381263	172699	553962
4. MEDAK	61512	22659	84171	2843	2864	2849	174884	64895	239779
5. HYDERABAD	0	0	0	0	0	0	0	0	0
6. RANGAREDDY	17116	7167	24283	2594	2373	2529	44399	17007	61406
7. MAHABOOBNAGAR	66075	30203	96278	2262	2171	2233	149451	65571	215022
8. NALGONDA	127904	84826	212730	3051	2714	2917	390235	230218	620453
9. WARANGAL	127156	53917	181073	3165	3163	3164	402403	170539	572942
10. KHAMMAM	109579	12143	121722	3076	2659	3034	337054	32288	369342
TELANGANA	**748713**	**297475**	**1046188**	**2934**	**2858**	**2913**	**2197030**	**850259**	3047289


AREA
350000
300000
250000
200000
150000
100000
50000
0
1 . ADILABAD
2 . NIZAMABAD
3 . KARIMNAGAR
4 . MEDAK
5 . HYDERABAD
6 . RANGAREDDY
7 . MAHABOOBNAGAR
8 . NALGONDA
9 . WARANGAL
10. KHAMMAM
AREA KHARIF
AREA RABI
AREA TOTAL


YIELD
4000
3500
3000
2500
2000
1500
1000
500
0
1 . ADILABAD
2 . NIZAMABAD
3 . KARIMNAGAR
4 . MEDAK
5 . HYDERABAD
6 . RANGAREDDY
7 . MAHABOOBNAGAR
8 . NALGONDA
9 . WARANGAL
10. KHAMMAM
YIELD KHARIF
YIELD KHARIF
YIELD TOTAL

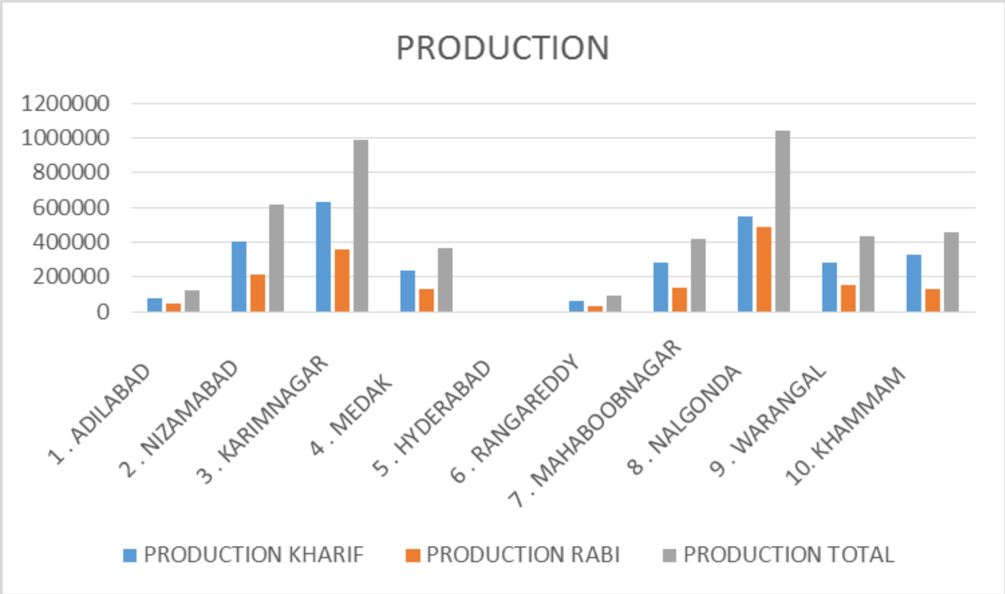


Figure 2: Area, Yield And Production Report For Rice Crop

VI. CONCLUSION

Agriculture is backbone to every country in the world either directly and indirectly. Mankind continues its journey with food produced by means of agriculture. From the review of literature it is understood that the state-of-the-art technology usage in agriculture has been disappointing. The adaptation of technology towards better productivity in agriculture has been slow and further research is needed to expedite it towards precision agriculture. Crop yield prediction is a significant component of national food security assessment and food policy making. Crop growth and yield data are critical for regulating agricultural cultivation system, and agricultural operation and management. The proposed crop yield prediction consists of three phases namely, preprocessing, feature reduction and prediction. Here the proposed method use input data as real world data. Real world data is often incomplete, inconsistent, and/or lacking in certain behaviors or trends, and is likely to contain many errors. Agriculture in the country can be improved with automatic prediction of crop yield based on the reliable variables. The system helps farmers to do right things at right time. The productivity gets improved in agriculture with sustained research in the field of spatial data mining to realize precision agriculture.

REFERENCE

[1] Ramesh A.Medar and Vijay.S.Rajpurohit, "A survey on Data Mining Techniques for Crop Yield Prediction", international journal of advance research in computer science and management studies, Vol.2, Issue.9, pp. 59-64, Sep 2014.

[2] M.C.S.Geetha, "A Survey on Data Mining Techniques in Agriculture", International journal of innovative research in computer and communication engineering, Vol.3, Issue.2, pp. 887-892, Feb 2015.
[3] Ch.N.Santhosh Kumar, V. Sitha Ramulu, K.Sudheer Reddy, Suresh Kotha and Ch. Mohan Kumar, "Spatial Data Mining using Cluster Analysis", International journal of computer science and information technology, Vol.4, No.4, pp. 71-77, Aug 2012.
[4] Raorane A.A. and Kulkarni R.V., "Review- Role of Data Mining in Agriculture", International journal of computer science and information technologies, vol.4, No.2, pp. 270-272, 2013.
[5] Bhavesh Kataria, "Weather-Climate Forecasting System for Early Warning in Crop Protection", Journal of engineering and technology, Vol.1, Issue.5, pp. 442-444, Sep 2015.
[6] He Li, Zhongxin Chen, Wenbin Wu, Zhiwei Jiang, bin Liu and Tuya Hasi, "Crop Model Data Assimilation with Particle Filter for Yield Prediction Using Leaf Area Index of Different Temporal Scales", International conference on agro geo informatics, pp. 401-406, July 2015.
[7] JU Hui, LIN Er-da, Tim Wheeler, Andrew Challinor and JIANG Shuai, "Climate Change Modeling and Its Roles to Chinese Crops Yield", Journal of integrative agriculture, Vol.12, No.5, pp. 892-902, 2013.
[8] SHI Wenjiao, TAO Fulu and ZHANG Zhao, "A review on statistical models for identifying climate contributions to crop yields", Journal of geographical sciences, Vol.23, No.3, pp. 567-576, 2013.
[9] M. Gunasundari, Dr. T. Arunkumar and Ms. R. Hemavathy, "CRY –An improved Crop Yield Prediction model using Bee Hive Clustering Approach for Agricultural data sets", International conference on pattern recognition, informatics and mobile engineering, pp. 473-478, 2013.
[10] Narendra Kumar Gontia and Kamlesh N.Tiwari, "Yield Estimation Model and Water productivity of Wheat crop in an Irrigation command using Remote Sensing and GIS", Journal of Indian society of remote sensing, Vol.39, No.1, pp. 27-37, Feb 2011.
[11] Manickam Gopperundevi and Vijayaraghavan Kannan, "Paddy Yield Estimation Using Remote Sensing and Geographical Information System", Journal of modern biotechnology, Vol.1, No.1, pp. 26-30, Sep 2012.
[12] E.I. Papageorgiou, A.T. Markinos and T.A. Gemtos, "Fuzzy cognitive map based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application", Vol.11, pp. 3643-3657, 2011.
[13] Xiaoqin Dai, Zailin Huo and Huimin Wang, "Simulation for response of crop yield to soil moisture and salinity with artificial neural network", International journal of field crops research, Vol.121, pp. 441-449, 2011.
[14] Luke Bornn and James V. Zidek, "Efficient stabilization of crop yield prediction in the Canadian Prairies", International journal of agricultural and forest meteorology, Vol.152, pp. 223-232, 2012.

[15] Dimitrios Voloudakis, Andreas Karamanos, Garifalia Economou, Dionissios Kalivas, Petros Vahamidis, Vasilios Kotoulas, John Kapsomenakis and Christos Zerefos, "Prediction of climate change impacts on cotton yields in Greece under eight climatic models using the AquaCrop crop simulation model and discriminant function analysis." Agricultural and water Management, Vol.147, pp. 116-128, 2015.

[16] David Gouache, Anne-Sophie Bouchon, Elodie Jouanneau and Xavier Le Bris, "Agro meteorological analysis and prediction of wheat yield at the departmental level in France", International journal of agricultural and forest meteorology, Vol.209, pp. 1-10, 2015.