International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2019 IJSRCSEIT | Volume 5 | Issue 1 | ISSN : 2456-3307
DOI : https://doi.org/10.32628/CSEIT1951122

A Survey on Rice Crop Yield Prediction in India Using Improved Classification Technique

Kolin Sukhadia*[1], M. B. Chaudhari[2]

[*1]PG Scholar, Department of Computer Engineering, Government Engineering College Modasa, Modasa, Gujarat, India
[2]Head of Department, Department of Computer Engineering, Government Engineering College Modasa, Modasa, Gujarat, India

ABSTRACT

India is an agricultural country. Agriculture is the important contributor to the Indian economy. There are many classification techniques like Support Vector Machine(SVM), LADTree, Naïve Bayes, Bayesnet, K Nearest Neighbour(KNN), Locally Weighted Learning(LWL) on rice crop production datasets. They have some drawbacks like low accuracy and more errors. To achieve more significant result, To increase classification accuracy and reducing classification errors, our research uses classification method Bayesnet based adaboost will be proposed in work. Rice crop yield depend on environment's parameters like Rainfall, minimum temperature, average temperature, Maximum temperature, Vapour Pressure, potential evapotranspiration, reference crop evapotranspiration, cloud cover, wet day frequency for the kharif season. our dataset containing these environmental parameters for accurate prediction of Rice crop yield.

Keywords : AdaBoost Technique, BayesNet Classifier, Rice Crop Yield Prediction, Kharif Season

I. INTRODUCTION

Rice crop production contributes more than 40% to overall crop production. Its production is reliant on climatic conditions. Improving the ability to predict crop productivity in under different climatic parameters like rainfall, temperature, cloud cover, vapour pressure, potential evapotranspirsation etc. can assist farmers and other stakeholders in making important decisions in terms of agronomy and crop choice[43]. The prediction of agricultural yield is a challenging and desirable task for every nation. Prediction of yield can be done based on historic crop cultivation data and environment data and environment data crop yield can be predicted. Crop yield prediction is the forecasting of harvest in Kg per Hectare. To analyse data various data mining technique can be used. Crop yield prediction can be used by Government, policy makers, agro-based industries, traders and agriculturists. The Government use crop yield prediction in procurement, distribution, buffer-stocking, import export, price fixation and marketing of agricultural commodities[8]. Agriculture is the most important branch in Indian economy and 70 percentage of rural population livelihood depends on agricultural work. Crop yield depends on environment's factors like precipitation, temperature, evapotranspiration, etc. Based on previous experience, Farmers cultivate crop. But nowadays, the uncertainty increased in environment. So, accurate analysis of historic data of environment parameters should be done for successful farming. To get more harvest, we should also do the analysis of previous cultivation data. The Prediction of crop yield can be done based on historic

crop cultivation data and weather data using data mining methods[14].

There are many classification techniques like Support Vector Machine(SVM), Naïve Bayes, BayesNet, K NearestNeighbour (KNN), Locally Weighted Learning(LWL), LAD Tree are applied on rice crop production datasets. The provision of accurate and timely information can assist farmers to make the best decision for their cropping situations This could benefit them to attain greater crop productivity[42]. This study examines the application of machine learning for the prediction of rice crop yield. The current study use the most rice producing states(Haryana, Punjab, Tamilnadu, Andhra Pradesh, Bihar). The dataset parameters are state name, district name, crop year, Rainfall, minimum temperature, average temperature, maximum temperature, vapour pressure, reference opevapo transpiration,
Potential evapotranspiration, diurnal temperature range, area, production, yield, class. The dataset was then sorted on the basis of yield to classify the records in to low, moderate and high. The low yield was from 0. 1 to 1 tonnes/hectare, moderate from 1 to 2 tonnes/hectare and high from 2 to 5
tonnes/hectare.

The further document is as follows. In Section 2 (Literature Survey), the literature review done on Rice Crop Yield Prediction, Bayesnet Method, AdaBoost method and Rice Crop. In Section 3 (Conclusion) conclusion derived based on literature survey.

II. LITERATURE SURVEY

A. CROP YIELD PREDICTION

IN [1] Umid Kumar Dey, Abdullah Hasan Masud, Mohammed Nazim Uddin does research to predict the yield of rice in the regions of Bangladesh during the aforementioned seasons using Multiple Linear Regression AdaBoost (Adaptive Boosting), Support Vector Machine Regression and a Modified Nonlinear Regression (MNR) equation and then comparing the modified equation with the other three methods to check its accuracy For Aus and Boro Rice our modified Nonlinear Regression equation not only yields better RMSE, MSE and MAE values, it also has the highest Rsquare value, which proves that the MNR equation is the best fit for this study. this study also proves that weather conditions play very vital role in predicting the yield. In [2], author used classifiers for research are J48, LWL, LAD Tree and IBK. The experimental results showed that J48, LADTree achieved the highest accuracy, sensitivity and specificity. In [3], author used classifiers were BayesNet and NaiveBayes. The results showed that the performance of BayesNet was much better compared with NaiveBayes for the dataset. In [4], proposed approach builds a model to predict the weather parameters applying Neural Networks, then estimates the rice yields applying Support Vector Regression. In [5], The authors have used pre-processed data that was clustered using k-means clustering algorithm. The association rule mining process will apply on clustered data to find the rules. The training phase ends with number of generated rules. In the testing phase, the yield value is predicted based on the generated rules. The parameters considered for research were precipitation, minimum temperature, reference crop evapotranspiration average temperature, maximum temperature. production, area and yield for the Kharif season including months from June to November.

This research discusses the experimental results obtained by applying SMO classifier using the WEKA tool on the dataset of 27 districts of Maharashtra state, India in [6]. In [7], The parameters considered for the present study were precipitation, reference crop evapotranspiration average temperature, minimum temperature, maximum temperature, area, production and yield for the Kharifs season (June to

November) for the years 1998 to 2002. In this research WEKA tool is used. A Multilayer Perceptron Neural Network was implemented as ANN. Cross validation method was used to validate data. In [8], The proposed method uses yield and weather data collected from United States Department of Agriculture. The parameters in the dataset are rainfall, humidity, temperature, yield and. regression-based algorithms to predict the crop yield. These were multivariate polynomial regression, support vector machine regression and random forest. support vector machine regression to obtain the best possible results for predicting the crop yield. In [9], The prominent classification methods in this study are the Support VectorMachine(SVM) and Naive Bayes. In this paper, AdaSVM and AdaNaive are the proposed ensemble model use to project the crop production over a period of time. This ensemble model is compared with Naive Bayes and SVM methods. accuracy and the classification error are two parameters used for output predictions. The finding yields that AdaSVM and AdaNaive are agreeable than SVM and Naive Bayes for the data set analyzed. In this study[10], regression based adaptive boosting prediction model is presented, There are three coastal districts belonging to Odisha are located in India using the datasets of Rabi and Kharif seasons. This study experiments on weak regressors, such as: SVR regression, linear, lasso, ridge proposes strong predictors by avoiding the shortcomingss of individual weak regressors. In [11] study examines the variations of climate variables and rice yield and examines the relationships using multiple linear regression, principal component analysis, and support vector machine (SVM) analysis in southwest Nigeria. In this study[12], Adaptive boosting of weak regressors for forecasting of crop production considering climatic variability. In this study[13] examines the variations of climate variables and rice yield and quantifies the relationships using multiple linear regression, principal component analysis, and support vector machine (SVM) analysis in southwest Nigeria. The climate data used was for a period of 36 years from 1980 to 2015. This study[14] describes the role of data mining in Agriculture and crop yield prediction. This paper also describes Groundnut crop yield prediction analysis and Naive Bayes Method. This Study [15] describes multi-model heterogeneous ensemble approach. This study [16] examine the introduction about Bayesian network, Directed acyclic graph, Conditional Probability table. This study [17] examine various data mining classification, clustering and association rule mining technique. This study [18] examine detail information about naïve bayes classification technique. This study [19] examine Bayesian network classification technique, structure learning technique, parameter learning technique. This study[20] examine various adaptive Learning Algorithms for Bayesian Network Classifiers. This study[21] examine Bayesian Network Classifiers in Weka. This study[22] examine the introduction about bayesian approach to Probability and Statistics. This study[23] examine bayesian method for constructing probabilistic networks from databases. they focus on constructing Bayesian belief networks. This study[24] examine the overview of rice crop yield prediction. Examines Different data mining techniques utilized for foreseeing rice crop yield. In this study[25], there is overview of rice crop yield prediction. This paper Examines Different data mining techniques utilized for foreseeing rice crop yield. In this study[26], rice yield prediction is performed by back propagation algorithm. In this study[27], there is Comparative Study of Classification Algorithms for Forecasting Rainfall. This study[28] examine Modeling Rainfall Prediction Using Data Mining Method: A Bayesian Approach data mining method for modeling rainfall prediction. In [29], Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India using data mining Techniques is applied. For data set results show that J48 and LADTree achieved the highest accuracy, sensitivity and specificity. Classification by LWL classifier displayed the lowest accuracy, sensitivity

and specificity results. In [30], This study presents an analysis using data mining techniques for estimating the future yield prediction in tea cultivation with climatic change trends observed in last 30 years (1977-2006). The patterns of crop production in response to the climatic (rainfall, temperature, relative humidity, evaporation and sunshine) effect across the four tea growing regions (South Bank, North Bank, Upper Assam and Cachar) of Assam were developed using Multiple Linear Regression (MLR) technique.

The tea production estimation equations developed for the regions were validated for the future yield prediction (2007, 2009 and 2010) and were found to be significant.

A study by [31] has done Crop Yield Prediction of Wheat Using Fuzzy C Means Clustering and Neural Network. A research work [32] has implemented artificial neural networks (ANNs) for predicting crop yields for an agricultural region in Nepal. In research, Agricultural data was collected from paddy field cultivation in the Siraha district, an eastern region in Nepal. climatic parameters are rainfall, maximum temperature and minimum temperature, fertilizer use were also used as input values. The result shows that the trained neural network produced a minimum error. In [33] researcher has implemented neural networks for rice yield prediction in mountainous regions. A study by [34] reaching review of research devoted to applications of machine learning in agrarian generation frameworks. The review reports the application of machine learning techniques with artificial neural networks, Bayesian networks, support vector machines and k-means. In [35], researchers predict Rice Production in Phimai District, Thailand Decision Support System Using Artificial Neural Network. A study by [36] implemented C 4. 5 Algorithm to predict yield of crops like paddy, maize, wheat, soybean. Interactive web pages developed for crop yield prediction. Accuracy is highest in case of Soybean. A study by [37] has applied association rules for Tamilnadu. Rice crop yield predicted using regression analysis, experiment implemented in MATLAB tool. Yield predicted from Annual Rainfall, Area under Cultivation, and Food Price Index. Relationship established between AR, AUC and FPI[38]. A study [39] described literature review in tabular form. Some research study only summarized the data mining techniques for crop yield prediction. Support Vector Machine, Random Forest, Neural Network, REPTree, Bagging, Bayes algorithm described in the study [40]. These classifiers used for Soybean Crop Yield Prediction and results are compared. Bagging algorithm gives maximum accuracy for soybean crop yield prediction among REPTree, Random Forest, Neural Network, SMO classifiers. Bayesian modelling used for wheat crop yield prediction using small dataset[41]. A study also identified drought related agronomic factors that significantly affect the variation in the rain fed wheat yield and relative importance of these factors.

Applications

Predicts crop yield production of rice for most Rice producing states, India.

Predicting productivity of crop in various climatic conditions can help farmer and other partners in essential basic leadership as far as agronomy and product decision. It also helps to increase the yield of crop.

III. PERFORMANCE EVALUATION METRICS

Accuracy: It is defined as the overall success rate of the classifier.

Accuracy = (TP + TN) / (TP + FP + FN + TN).

Where, True Positive (TP) depicts the number of instances where system detects for condition when it is really present[1].

True Negative (TN) depicts the number of instances where system does not detect a condition when it is absent[1].

False Negative (FN) depicts the number of instances where system does not detect a condition when actually it is present[1].

False Positive (FP) depicts the number of instances where system detects a condition when it is really absent[1].

RMSE(Root Mean Square Error): It is defined as the difference between the values predicted by the model and the actual values noted[1].

MAE(Mean Absolute Error): It is another factor in statistics which measures the difference between two continuous variables[1].

RAE(Relative Absolute Error): This measure gives the total absolute error between the variables[1].

IV. CONCLUSION AND FUTURE SCOPE

In rice crop yield prediction, various classification techniques are applied like support vector machine, naïve bayes, bayesnet, locally weighted learning, ladtree etc. are applied on rice crop production datasets. They have some drawbacks of low classification accuracy and high errors. To get more significant result and to get better classification accuracy and to reduce errors like RMSE(Root Mean Square Error), MAE(Mean Absolute Error), RAE(Relative Absolute Error), boosting technique Adaptive boosting technique is used in conjuction with classification method Bayesnet in my research work and that will determine better accuracy of prediction of rice yield in future. This survey concludes that the more accurate rice crop yield prediction model can be constructed by using different data mining techniques.

V. REFERENCES

[1]. Shruti Mishra, Priyanka Paygude, Snehal Chaudhary, Sonali Idate"Use of Data Mining in Crop Yield Prediction", Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018), pp. 796-802, 2018

[2]. Umid Ku mar Dey, Abdullah Hasan Masud, Mohammed Nazim Uddin, "Rice Yield Prediction Model Using Data Mining", International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 321-326, 2017

[3]. Niketa Gandhi, Leisa J. Armstrong, Manisha Nandawadekar, "Application Of Data mining Techniques For Predicting Rice Crop Yield Semi-Arid Climate Zone Of India", 2017 IEEE International Conference on Technological Innovations in ICT for agriculture and Rural Development, pp. 116-120, 2017

[4]. Ayush Shah, Akash Dubey, Vishesh Hemnani, Divye Gala, D. R. Kalbande, "Smart Farming System: Crop Yield Prediction Using Regression Techniques", Proceedings of International Conference on Wireless Communication, pp. 49-56, 2018

[5]. Niketa Gandhi, Leisa J. Armstrong, Owaiz Petkar, Amiya Kumar Tripathy, "Rice Crop Yield Prediction in India using Support Vector Machines", 2016 International Joint Conference on Computer Science and Software Engineering (JCSSE)

[6]. Niketa Gandhi, Leisa J. Armstrong, Owaiz Petkar, "Predicting Rice Crop Yield Using Bayesian Networks", 2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 795-799, 2016

[7]. Akter Hossain, Mohammed Nazim Uddin, Mohammad Arif Hossain, Yeong Min Jang, "Predicting Rice Yield For Bangladesh By Exploiting Weather Condition", 2017 International Conference on Information and

Communication Technology Convergence (ICTC), pp.589-594, 2017

[8]. Dipika H. Zala, M.B. Chaudhari, " Review on Use of "BAGGING" Technique in Agriculture Crop Yield Prediction", International Journal for Scientific Research & Development| Vol. 6, Issue 08, 2018, pp.675-677.

[9]. Niketa Gandhi, Leisa J. Armstrong, "Applying Data Mining Techniques to Predict Yield Of Rice in Humid Subtropical Climatic Zone Of India", 2016 International Conference on Computing for Sustainable Global Development, pp.1901-1906, 2016

[10]. Sagar S. Nikam , "A Comparative Study of Classification Techniques in Data Mining Algorithms", ISSN April 2015.

[11]. Narayanan Balakrishnan, Dr.Govindarajan Muthukumarasamy, "crop Production - Ensemble Machine Learning Model for Prediction", International Journal of Computer Science and Software Engineering (IJCSSE), Volume 5, Issue 7, July 2016, pp. 148-153.

[12]. Subhadra Mishra, Debahuti Mishra, Gour Hari Santra, "Adaptive boosting of weak regressors for forecasting of crop production considering climatic variability: An empirical assessment", Journal of King Saud University - Computer and Information Sciences, pp. 1-23.

[13]. Philip G. Oguntunde, Gunnar Lischeid, Ottfried Dietrich, "Relationship between rice yield and climate variables in southwest Nigeria using multiple linear regression and support vector machine analysis", August 2017, springer.

[14]. Haresh L. Siju, Pinal J. Patel, "Review on Crop Yield Prediction using Data Mining Focusing on Groundnut Crop and Naive Bayes Technique", International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Volume 3 | Issue 1, ISSN : 2456-3307, pp.1947-1954

[15]. Olugbenga Wilson Adejo, Thomas Connolly, "Predicting student academic performance using multi-model heterogeneous ensemble approach", Journal of Applied Research in Higher Education, Vol. 10 Issue: 1, pp.61-75

[16]. ADNAN MASOOD, "Bayesian Networks", A BRIEF INTRODUCTION

[17]. Jiawei Han, Micheline Kamber, "Data Mining:Concepts and Techniques"

[18]. https://en.wikipedia.org/wiki/Naive_Bayes_classifier

[19]. Friedman, N., Geiger, D. and Goldszmidt M. (1997), " Bayesian network classifiers", Machine Learning 29, 139–164.

[20]. Castillo, G. (2006) , "Adaptive Learning Algorithms for Bayesian Network Classifiers", PhD Thesis, Chapter 3.

[21]. Bouckaert, R. (2007), " Bayesian Network Classifiers in Weka", Technical Report

[22]. Heckerman, D. (1996), " A Tutorial on Learning with Bayesian Networks", Microsoft Technical Report MSR-TR-95-06.

[23]. G. Cooper, E. Herskovits (1992), "A Bayesian method for the induction of probabilistic networks from data, "Machine Learning. 9(4):309-347.

[24]. N.Gandhi, L.J. Armstrong and O. Petkar, "Rice Crop Yield Prediction in India using Machine Learning Techniques", communicated, 2016.

[25]. Weka 3:Data Mining Software in Java, Machine Learning Group at the University of Waikato, Official Web:http://www.cs.waikato.ac.nz/ml/weka/index.html, accessed on 26th March 2016.

[26]. S. Puteh, M. Rizon, M. Juhari, J. Nor Khairah, S. Siti Kamarudin, B. Aryati, R. Nursalasawati, "Back propagation algorithm for rice yield prediction", Proceedings of 9th of the Ninth International Symposium on Artificial Life and Robotics, Beppu, Japan, Oita, pp. 586-589, 2004.

[27]. Deepti Gupta, Udayan Ghose, " A Comparative Study of Classification Algorithms for Forecasting Rainfall ", IEEE 4th International conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), 2015, pp. 1-6.

[28]. V. B. Nikam and B. B. Meshram, "Modeling Rainfall Prediction Using Data Mining Method: A Bayesian Approach", Fifth International Conference on Computational Intelligence, Modelling and Simulation, Seoul, 2013, pp. 132-136.

[29]. Niketa Gandhi, Leisa J. Armstrong, "Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India using data mining Techniques", 2016 IEEE International Conference on Advances in Computer Applications (ICACA)

[30]. Rupanjali D. Baruah, R.M. Bhagat, Sudipta Roy, L.N. Sethi, "Use of Data Mining Technique for Prediction of Tea Yield in the Face of Climate Change of Assam, India", 2016 International Conference on Information Technology, pp. 265-269.

[31]. Akanksha Verma, Aman Jatain, Shalini Bajaj, " Crop Yield Prediction of Wheat Using Fuzzy C Means Clustering and Neural Network", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 11 (2018) pp. 9816-9821.

[32]. Tirtha Ranjeet, Leisa Armstrong, " An Artificial Neural Network for Predicting Crops Yield in Nepal", Proceedings of AFITA 2014 - Copyright ASICTA Inc, pp. 376-386.

[33]. B. J I , Y. SUN, S.YANG, J. WAN1, " Artificial neural networks for rice yield prediction in mountainous regions", Journal of Agricultural Science (2007), pp. 249–261.

[34]. Ankita Bissa, Meena Kushwaha , Mayank Patel, "A Survey of Machine Learning Appliacations In Decision Making To Improve Farming", International Journal of Computer Sciences and Engineering, Vol.-6, Issue-9, Sept. 2018, pp. 669-683.

[35]. Saisunee Jabjone, Sura Wannasang, " Decision Support System Using Artificial Neural Network to Predict Rice Production in Phimai District, Thailand", International Journal of Computer and Electrical Engineering, Vol. 6, No. 2, April 2014, pp. 162-166.

[36]. Veenadhari, B. Misra, and C. D. Singh, "Machine learning approach for forecasting crop yield based on climatic parameters, " 2014 Int. Conf. Comput. Commun. Informatics Ushering Technol. Tomorrow, Today, ICCCI 2014, pp. 1-5, 2014.

[37]. E. Manjula, S. Djodiltachoumy, " A Model for Prediction of Crop Yield", International Journal of Computational Intelligence and Informatics, Vol. 6: No. 4, March 2017, pp.298-305.

[38]. Sellam and E. Poovammal, "Prediction of Crop Yield using Regression Analysis, " vol. 9, no. October, 2016.

[39]. B. H. Dhivya, R. Manjula, S. B. S, and R. Madhumathi, "A Survey on Crop Yield Prediction based on Agricultural Data, " pp. 4177-4183, 2017.

[40]. Savla and A. Mandholia, "Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture, " 2015.

[41]. Yildirak, Z. Kalaylioglu, and A. Mermer, "Bayesian estimation of crop yield function : drought based wheat prediction model for tigem farms, " Environ. Ecol. Stat., vol. 22, no. 4, pp. 693-704, 2015.

[42]. Mrs. K. E. Eswari, Ms. L.Vinitha, " Crop Yield Prediction in Tamil Nadu using Baysian Network", International Journal of Intellectual Advancements and Research in Engineering Computations, Volume-6, Issue-2, pp.1572-1576.

[43]. Niketa Gandhi, Owaiz Petkar, Leisa J. Armstrong, "Rice Crop Yield Prediction Using Artificial Neural Networks", 2016 IEEE International Conference on Technological Innovations in ICT For Agriculture and Rural Development (TIAR 2016), pp.105-110

Cite this article as : Kolin Sukhadia, M. B. Chaudhari, "A Survey on Rice Crop Yield Prediction in India Using Improved Classification Technique ", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 5 Issue 1, pp. 501-507, January-February 2019.

Available at doi : https://doi.org/10.32628/CSEIT1951122

Journal URL : http://ijsrcseit.com/CSEIT1951122