IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Crop Yield Prediction and Efficient use of Fertilizers

Aruna Kamble
(Faculty)
Department of CSE
Bharati Vidyapeeth
College of Engineering
Belapur, Navi Mumbai
Maharashtra, India

Vibhoor Garg
(Student)
Department of CSE
Bharati Vidyapeeth
College of Engineering
Belapur, Navi Mumbai
Maharashtra, India

Abstract—In India, all of us understand that Agriculture is the spine of the country. This paper predicts the yield of virtually all sorts of vegetation which can be planted in India. This script makes novel via way of means of using simple parameters like State, district, season, place and the user can predict the yield of the crop in which year she or he wishes to. The paper makes use of superior regression techniques like Kernel Ridge, Lasso and ENet algorithms to expect the yield and makes use of the concept of Stacking Regression for reinforcing the algorithms to offer a better prediction.

***Keywords*— Crop yield prediction, Lasso, Kernel Ridge, ENet, Stacked Regression.**

INTRODUCTION

In our studies, which we discovered in the preceding studies papers is that everybody makes use of climatic elements like rainfall, daylight and agricultural elements like soil type, nutrients possessed in the soil (Nitrogen, Potassium, etc.) however the hassle is we want to acquire the data after which a third party does this prediction and then it is told to the farmer and this takes numerous attempts for the farmer and he doesn't apprehend the technology in the back of those factors. To make it easy and which can be immediately utilized by the farmer this paper uses simple elements like which state and district is the farmer from, which crop and in what season (as in Kharif, Rabi, etc.). In India, there are more than one hundred crops planted across the entire country. These vegetations are classified for higher understanding and visualization. The data for this study has been obtained from the Indian Government Repository [1]. The data includes attributes – State, District, Crop, Season, Year, Area and Production with around 2.5 Lakh observations. The fig. 1. depicts the states and territories of India which visualize that which class of crops are well-known in which season. We used superior regression strategies – Lasso, ENet and Kernel Ridge and similarly we used stacking of those models to decrease the error and to achieve better predictions. This paper set out as follows: Literature Survey, Methodology, Conclusion and Future Work.

LITERATURE SURVEY

Ananthara, M. G. et al. (2013, February) proposed a prediction model for datasets bearing on agriculture that's referred to as CRY algorithm for crop yield by using beehive clustering techniques. They have taken into consideration, the parameters particularly crop kind, soil type, soil pH value, humidity and crop sensitivity. Their analysis was in particular in paddy, rice and sugarcane yields in India. Their proposed algorithm was later then compared with C&R tree algorithm and it outperformed nicely with an accuracy of ninety percent [2]. Awan, A. M. et al. (2006, April) constructed a brand new, clever framework targeted on farm yield prediction clustering kernel method and in that, they have taken into consideration the parameters like plantation, latitude, temperature and precipitation of rainfall in that latitude. They had experimented weighted k-means kernel approach with spatial constraints for the evaluation of oil palm fields [3]. Chawla, I. et al. (2019, August) used fuzzy logic for crop yield prediction via statistical time series models.

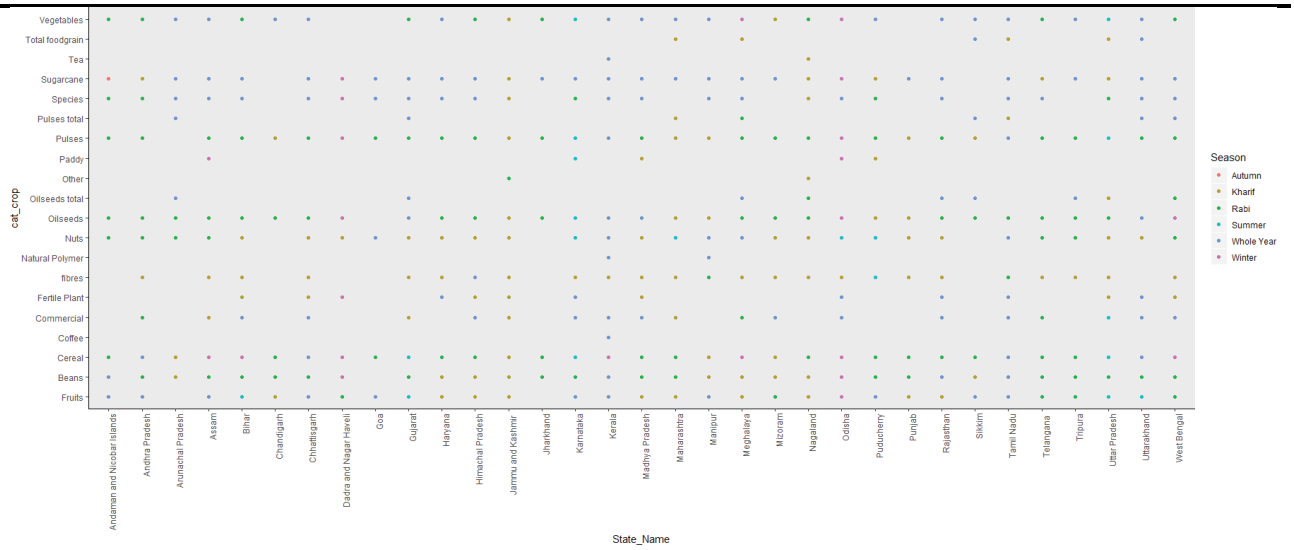

Fig. 1. Famous Categories of crops over states in India (based on Season)

They considered parameters like rainfall and temperature for prediction. Their prediction turned into classification with levels 'good yield', 'very good yield' [4]. Chaudhari, A. N. et al. (2018, August) used 3 algorithms particularly clustering k- manner, Apriori and Bayes algorithm, then they hybridized the algorithm for higher efficiency of yield prediction and then they considered parameters like Area, Rainfall, Soil type and additionally their system was able to inform which crop is appropriate for cultivation primarily based totally on the mentioned features [5]. Gandge, Y. (2017, December) used many machine learning algorithms for distinct crops. They studied and analyzed which algorithm might be appropriate for which crop. They have used K-means, Support vector Regression, Neural Networks, C4.5 Decision tree, Bee-Hive Clustering, etc. The elements implying had been soil nutrients like N, K, P and soil ph. [6]. Armstrong, L. J. et al. (2016, July) used ANNs for the prediction of rice yield all over the districts of Maharashtra, India. They had taken into consideration the climatic factors particularly (considering range) temperature, precipitation and reference crop evapotranspiration. The information has been accumulated from Indian Government repository from 1998 to 2002 [7]. Tripathy, A. K. et al. (2016, July) were same authors who used encourage the use of vector machines to predict the rice crop yield with equal functions as the preceding paper noted [8]. Petkar, O. (2016, July) was also the same authors who implemented for SVM and neural networks for rice crop yield prediction proposed a brand new decision system that is an interface to offer the input and get the output [9]. Chakrabarty, A. et al. (2018, December) analyzed crop prediction in the nation of Bangladesh in which they majorly cultivate 3 sorts of rice, Jute, Wheat, and Potato. Their studies used a deep neural network in which the data had round forty-six parameters into their consideration. Few of them have been soil composition, kind of fertilizer, type of soil and its structure, soil consistency, response and texture [10]. Jintrawet, A. et al. (2008, May) used SVR model for crops like rice to expect the yield in which the model became divided into 3 steps- predicting the soil nitrogen weight observed by prediction of rice stem weight and rice grain weight respectively. Their elements have been solar radiation, temperature and precipitation together with those 3 steps [11]. Miniappan,
N. et al. (2014, August) used synthetic neural network in modelling multi-layer perceptron model with 20 hidden layers for prediction wheat yield which considered elements like daylight, rain, frost and temperature [12]. Manjula, A et al. constructed a crop selection and to predict the yield which had considered a numerous indexes like vegetation, temperature and normalized distinction vegetation as factors. They differntiated among climate elements and agronomic elements and some other disturbances induced in the prediction for better expertise [13]. Mariappan, A. K. et al. analyzed the data concerning rice crop withinside the state of Tamil Nadu, India. They have considered the elements like soil, temperature, sunshine, rainfall, fertilizer, paddy, and kind of pest used and some other elements like pollutants and season [14]. Verma, A. et al. (2015, December) used classification strategies like Naïve Bayes, K-NN algorithm for crop prediction on soil datasets which constituted nutrients of soil like zinc, copper, manganese, pH, iron, Sulphur, Phosphorous, Potassium, nitrogen, and Organic Carbon [15]. Kalbande, D. R. et al. (2018) used support vector regression, multi polynomial regression and random forestland regression for prediction of corn yield and evaluated the models with the use of metrics like errors to be named, MAE, RMSE and R-rectangular values [16]. Rahman, R. M. et al. (2015, June) used in particular clustering techniques for crop yield prediction. The paper defined the analysis of most major crops in Bangladesh and divided the variables into environmental and biotic variables. The algorithms implemented were namely linear regression, ANN, and KNN method for classification [17]. Hegde, M. et al. (2015, June) used a couple of linear regression and neuro fuzzy structures for predicting crop yield via way of means of taking biomass, soil water, radiation and rainfall as input parameters for the studies and their majorly focused crop became wheat [18]. Sujatha, R., & Isakki, P. (2016, January) used classification techniques like ANN, j48, Naïve Bayes, Random Forest and Support vector Machines. They have additionally covered each climatic parameters and soil parameters as functions of their modelling [19]. Ramalatha, M. et al. (2018, October) used a hybrid method of mixing K- means clustering and classification primarily based totally on modified K-NN method. The information was accumulated from Tamil Nadu, India in which the majorly focused crops have been rice, maize, Ragi, Sugarcane, and Tapioca [20]. Singh, C. D. et al. (2014, January) developed an utility to endorse crops which fits on selected districts of Madhya Pradesh, India. The user might give input cloud cover, rainfall, temperature, observing yield they had in the past and the system might expect the yield and Depending on the triggering values set, the crop can be categorized and obtain the results [21].

METHODOLOGY

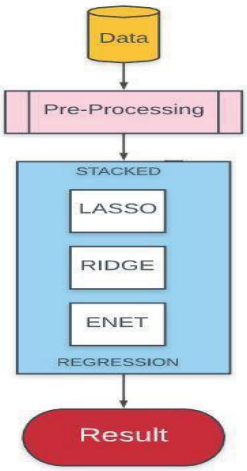

Fig. 2. Process chart of the research project

Pre-processing

For the given data set, there are many a few 'NA' values that are filtered in python. Furthermore, because the information set includes numeric information, we used sturdy scaling, that's pretty much like normalization, however it alternatively makes use of the interquartile range while normalization is something which normalization shrinks the data in terms of 0 to 1.

Stacked Regression:

This is a form of ensembling however a touch of enhancement of averaging. In this, we upload a meta model and use the out of fold predictions of the alternative models used to teach the main meta model.
Step-1: the total training set is once more divided into distinct sets. (Train and holdout)
Step-2: train the chosen base models with first part (train).
Step-3: Test them with the second of the parts. (holdout)
Step-4: Now, the predictions acquired from test component are inputs to the train better level learner referred to as meta-model.

Iteratively, the primary three steps are completed. For example, if we take a five-fold stacking, we divide the training information into five folds first. We'll then do five iterations. We train every base model on four folds in every iteration and predict the ultimate fold (holdout fold). So, after five iterations, we will be assured each one of the considered data can be used to get out - of-fold predictions that we will use as a brand new function in Step four to train our meta-model. We average up the predictions of all base models at the test data for the predictive part and used them as meta-functions on which the meta-model is eventually predicted. Here, our meta model is Lasso Regressor and that's the purpose for being placed on the pinnacle in fig. 2. The stacked regression functioning will be understood from the fig. 3.

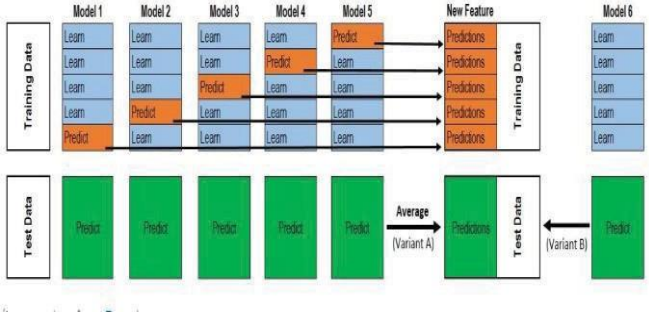

Fig. 3. Stacked Regression

CONCLUSION AND FUTURE WORK

The user or the farmer can input the subsequent information over the internet software to get the prediction. The output is presently an internet application, however our future workings might be constructing an application software in which the farmers can use it as app and changing the entire system into their regional language.

REFERENCES

[1] "data.gov.in." [Online]. Available: https://data.gov.in/

[2] Ananthara, M. G., Arunkumar, T., & Hemavathy, R. (2013, February). CRY—an improved crop yield prediction model using bee hive clustering approach for agricultural data sets. In *2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering* (pp. 473-478). IEEE.

[3] Awan, A. M., & Sap, M. N. M. (2006, April). An intelligent system based on kernel methods for crop yield prediction. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 841-846). Springer, Berlin, Heidelberg.

[4] Bang, S., Bishnoi, R., Chauhan, A. S., Dixit, A. K., & Chawla, I. (2019, August). Fuzzy Logic based Crop Yield Prediction using Temperature and Rainfall parameters predicted through ARMA, SARIMA, and ARMAX models. In *2019 Twelfth International Conference on Contemporary Computing (IC3)* (pp. 1-6). IEEE.

[5] Bhosale, S. V., Thombare, R. A., Dhemey, P. G., & Chaudhari, A. N. (2018, August). Crop Yield Prediction Using Data Analytics and Hybrid Approach. In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)* (pp. 1-5). IEEE.

[6] Gandge, Y. (2017, December). A study on various data mining techniques for crop yield prediction. In *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)* (pp. 420-423). IEEE.

[7] Gandhi, N., Petkar, O., & Armstrong, L. J. (2016, July). Rice crop yield prediction using artificial neural networks. In *2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)* (pp. 105-110). IEEE.

[8] Gandhi, N., Armstrong, L. J., Petkar, O., & Tripathy, A. K. (2016, July). Rice crop yield prediction in India using support vector machines. In 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 1-5). IEEE.

[9] Gandhi, N., Armstrong, L. J., & Petkar, O. (2016, July). Proposed decision support system (DSS) for Indian rice crop yield prediction. In 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR) (pp. 13-18). IEEE.

[10] Islam, T., Chisty, T. A., & Chakrabarty, A. (2018, December). A Deep Neural Network Approach for Crop Selection and Yield Prediction in Bangladesh. In 2018 IEEE Region 10 Humanitarian Technology Conference (R10-HTC) (pp. 1-6). IEEE.

[11] Jaikla, R., Auephanwiriyakul, S., & Jintrawet, A. (2008, May). Rice yield prediction using a support vector regression method. In 2008 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (Vol. 1, pp. 29-32). IEEE.

[12] Kadir, M. K. A., Ayob, M. Z., & Miniappan, N. (2014, August). Wheat yield prediction: Artificial neural network based approach. In 2014 4th International Conference on Engineering Technology and Technopreneuship (ICE2T) (pp. 161-165). IEEE.

[13] Manjula, A., & Narsimha, G. (2015, January). XCYPF: A flexible and extensible framework for agricultural Crop Yield Prediction. In 2015 IEEE 9th International Conference on Intelligent Systems and Control (ISCO) (pp. 1-5). IEEE.

[14] Mariappan, A. K., & Das, J. A. B. (2017, April). A paradigm for rice yield prediction in Tamilnadu. In 2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR) (pp. 18-21). IEEE.

[15] Paul, M., Vishwakarma, S. K., & Verma, A. (2015, December). Analysis of soil behaviour and prediction of crop yield using data mining approach. In 2015 International Conference on Computational Intelligence and Communication Networks (CICN) (pp. 766-771). IEEE.

[16] Shah, A., Dubey, A., Hemnani, V., Gala, D., & Kalbande, D. R. (2018). Smart Farming System: Crop Yield Prediction Using Regression Techniques. In Proceedings of International Conference on Wireless Communication (pp. 49-56). Springer, Singapore.

[17] Ahamed, A. M. S., Mahmood, N. T., Hossain, N., Kabir, M. T., Das, K., Rahman, F., & Rahman, R. M. (2015, June). Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh. In 2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD) (pp. 1-6). IEEE.

[18] Shastry, A., Sanjay, H. A., & Hegde, M. (2015, June). A parameter based ANFIS model for crop yield prediction. In 2015 IEEE International Advance Computing Conference (IACC) (pp. 253-257). IEEE.

[19] Sujatha, R., & Isakki, P. (2016, January). A study on crop yield forecasting using classification techniques. In 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16) (pp. 1-4). IEEE.

[20] Suresh, A., Kumar, P. G., & Ramalatha, M. (2018, October). Prediction of major crop yields of Tamilnadu using K-means and Modified KNN. In 2018 3rd International Conference on Communication and Electronics Syst ems (ICCES) (pp. 88-93). IEEE.

[21] Veenadhari, S., Misra, B., & Singh, C. D. (2014, January). Machine learning approach for forecasting crop yield based on climatic parameters. In 2014 International Conference on Computer Communication and Informatics (pp. 1-5). IEEE.