International Journal of Communication and Information Technology

E-ISSN: 2707-6628
P-ISSN: 2707-661X
IJCIT 2020; 1(2): 24-27
Received: 07-05-2020
Accepted: 10-06-2020

Punith Sai V
GATE College, Tirupati, Andhra Pradesh, India

Crop yield prediction for formers using random forest approach

Punith Sai V

DOI: https://doi.org/10.33545/2707661X.2020.v1.i2a.15

Abstract
As an agricultural region, India's economy depends primarily on agricultural yield growth and agro-industry goods. Data Mining is an new area of study in crop yield analysis. Prediction of yields is a very critical issue in agriculture. Any farmer is interested in knowing how much yield he's going to make. Analyze various related attributes such as location, pH, etc. Price from which the soil alkalinity is determined. In comparison, the percentage of nutrients such as Nitrogen (N), Phosphorous (P) and Potassium (K) Position is used along with the use of third-party applications such as environment and temperature APIs, soil quality, soil nutrient content in that area, amount of rainfall in the field, soil structure can be calculated. All of these data attributes will be analyzed, or the data will be developed with various correct machine learning algorithms to create a model. The system comes with a model to be accurate and accurate in predicting crop yields and to provide the end user with correct recommendations on the required fertilizer ratio based on the atmospheric and soil parameters of the soil to increase crop yields, and increase farmer revenue.

Keywords: Prediction, attributes, machine learning, algorithms, accuracy, crop yield

Introduction
Is a densely populated country and a random change in climate conditions is required to protect the world's food supplies. Framers face severe problems in dry conditions. The soil type plays a key role in crop yields. Suggesting the use of fertilizers can enable farmers to make the best decisions about their cropping situation. List of research on information and communication technology (ICT) could be used to estimate crop yield. Through using Data Mining, we can also estimate crop yields [1, 2, 3]. By thoroughly analyzing the previous data, we can recommend a farmer for a better crop for better yield The task is to create an efficient model to predict the most effective model to predict the performance of the crop. Try using different algorithms and compare all algorithms and which one has the least error and loss the model has selected and predict the yield of that particular crop. From this paper, u can see a comparison of the two algorithms and predict the output of the best model.

Related Works
Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone
Data mining is the process of identifying hidden patterns of large and complex data. It may play a crucial role in decision-making on complex agricultural concerns. Data visualization is equally important for understanding the general trends in the effect of the major factors influencing crop yields. The present study examines the application of data visualization techniques to the discovery of The study also applies data mining techniques to extract information from historical agricultural data set to predict the yield of crop production for the Kharif Season of the Tropical Wet and Dry Weather state of India [5, 6]. The data set were visualized in Microsoft Office Excel using scatter plots [5]. Classification algorithms have been executed in the WEKA free and open source data mining tool. The experimental results provided include sensitivity, specificity, accuracy, F1 score, Mathew coefficient of correlation, mean absolute error, root mean squared error, relative absolute error, and root relative squared error.

Crop Yield Prediction in Tamil Nadu Using Bayesian Network
Crop prediction is a large agricultural problem. In order to address this problem, clustering and classification techniques are used to predict crop yields. It is one of the most commonly

Corresponding Author:
Punith Sai V
GATE College, Tirupati, Andhra Pradesh, India

used expert systems based on data analytics principles to predict crop yield in order to maximize crop productivity [6, 16]. Machine learning techniques can be used to improve agricultural productivity prediction in different climate scenarios. Bayesian Network Classification is a supervised learning model, which means that temperature and rainfall evaluates the crop data used for Rice, Coconut, groundnut, Black Pepper and Dry Ginger crop classification and probability values.

Proposed Method

Algorithm

Random Forest

Random Forest is a very popular and powerful supervised machine learning algorithm that can perform both classification and regression tasks by building deterministic trees during class or predictive class training and performance. (Regression) of each tree. More trees in the park. They're stronger than expected. Random decision trees are ideally suited to the habit of a tree. Deciding trees that match their training package. The data sets considered are the array of decision trees that take into account rainfall, awareness, growth, random forest temperature, and two-thirds of the data sets. These tree decisions will be applied to the remaining records for accurate classification. The resulting training sets can be applied to the test data in a precise manner. The RF algorithm was used to study the performance of this methodology in the dataset. The advantage of stochastic forest algorithms is that the problem of overlapping with random forests is minimal, unlike decision tree machine learning algorithms.

Methodology

Step 1 − First, start with the selection of random sample form the given dataset.

Step 2 − Next, this algorithm will build a decision tree for each sample. At the point we will get the predicted result from each predicted results.

Step 3 − In this step, voting will be performed for every prediction result.

Step 4 − At last, select the most voted prediction result as the final result.

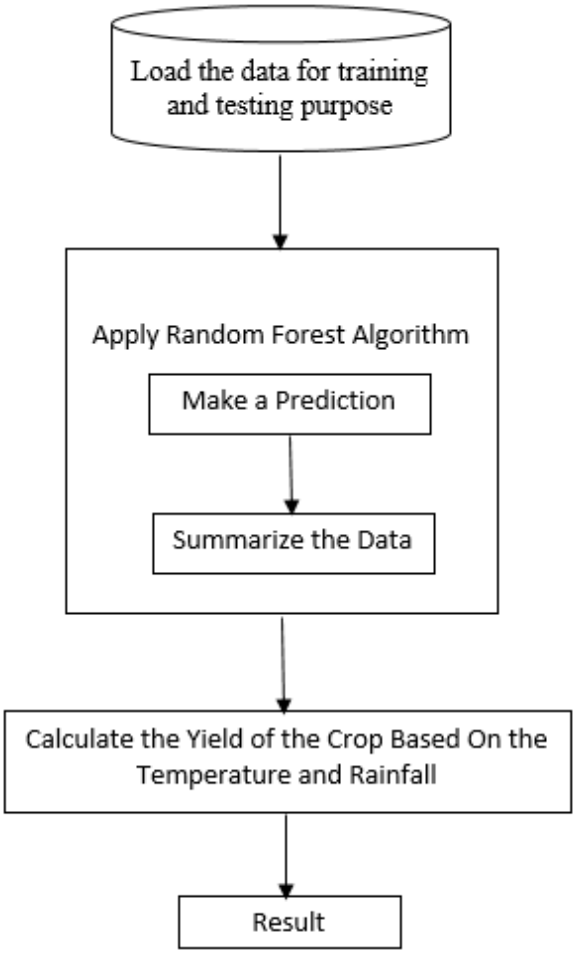

Fig 1: Flow of the model

Results and Discussion

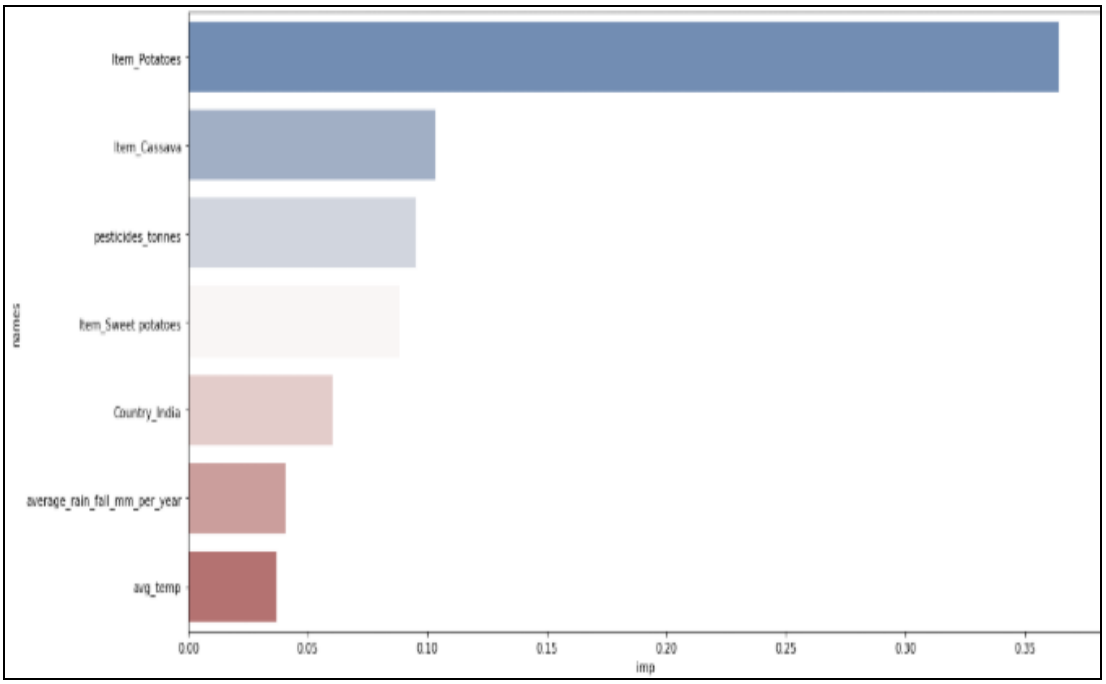

Fig 1: Feature Importance Bar Plot

Here in the above-mentioned Fig 1 it is representing the process of feature extraction of the given dataset. By using those features we will next be applying the prediction model and representing actual predictions in the Fig 2.

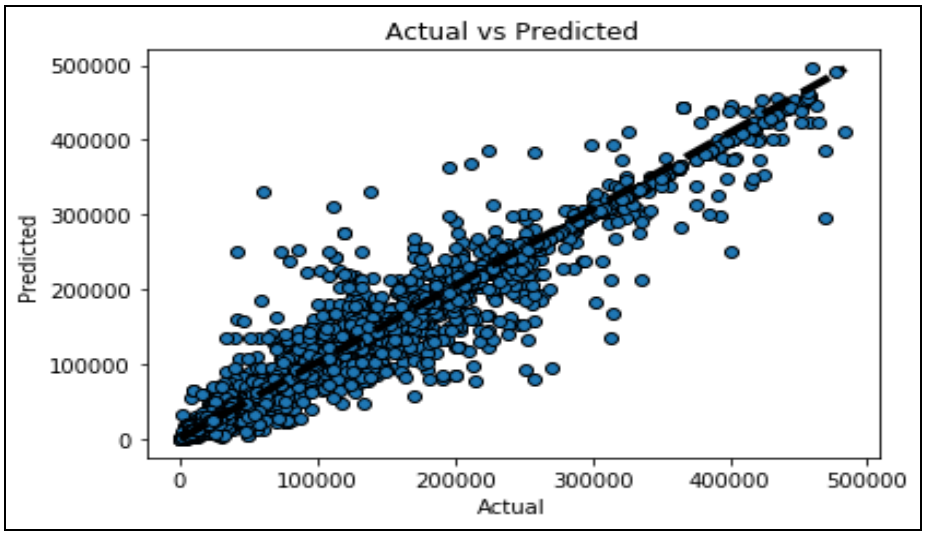

Fig 2: Actual Predict

The above Fig 2 graph representing the comparison of the actual and predicted values of the original data

Conclusion

Work crop yield prediction and efficient use of the fertilizer is successfully predicted and also found the efficient algorithm from both the algorithm and obtained the most efficient output of the yield. In future developing the web application based on this ideology and make the user use this easily and help the user to understand the yield of the crop, he is going to crop in that season the methodology utilized. The deep neural system to make yield prediction including yield, check yield and yield difference in view of genotype and condition information. We introduced a ML approach for crop yield prediction, which exhibits superior performance deliberately planned deep neural network had the option to learn nonlinear and complex connections between qualities, natural conditions, just as their associations from authentic information and make sensibly precise predict of yields for new hybrid plants in new location with weather conditions. The performance of the model was seen as moderately delicate to the nature of climate forecast, which proposed the significance of weather prediction strategies.

References

1. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J *et al*. TensorFlow: a framework for huge scope AI," in OSDI'16 Proceedings of the twelfth USENIX Conference on Operating Systems Design and

Implementation. 2016; 16(Savannah, GA):265-283.
2. Abhishek K, Singh M, Ghosh S, Anand A. Climate determining model utilizing fake neural system. Procedia Technol. 2012; 4:311-318. DOI: 10.1016/j.protcy.2012.05.047
3. Baboo SS, Shereef IK. An effective climate determining framework utilizing fake neural system. Int. J Environ. Sci. Dev. 2010; 1:321. DOI: 10.7763/IJESD.2010.V1.63
4. Bengio Y, Simard P, Frasconi P. Learning long haul conditions with slope drop is troublesome. IEEE Trans. Neural Netw. 1994; 5:157-166. DOI: 10.1109/72.279181
5. Bou-Rabee M, Sulaiman SA, Saleh MS, Marafi S. Utilizing fake neural systems to assess sun powered radiation in kuwait. Inexhaustible Sustain. Vitality Rev. 2017; 72:434-438. DOI: 10.1016/j.rser.2017.01.013
6. Burgueño J, Crossa J, Cornelius PL, Yang RC. Utilizing factor investigative models for joining conditions and genotypes without hybrid genotype × condition connection. Harvest Sci. 2008; 48:1291-1305. DOI: 10.2135/cropsci2007.11.0632
7. Mustafa AElRA, Moursy ARA. Using a multivariate regression model and hyperspectral reflectance data to predict soil parameters of Agra, India. Int. J Geogr Geol. Environ 2020;2(1):04-09. DOI: 10.22271/27067483.2020.v2.i1a.12
8. Bustami R, Bessaih N, Bong C, Suhaili S. Fake neural system for precipitation and water level forecasts of bedup waterway. IAENG Int. J. Comput. Sci. 2007; 34:228-233.
9. Chapman SC, Cooper M, Hammer GL, Butler DG. Genotype by condition communications influencing grain sorghum. II. Frequencies of various occasional examples of dry season pressure are identified with area consequences for crossover yields. Aust. J. Agric. Res. 2000; 51:209-222. DOI: 10.1071/AR99021
10. Cooper M, DeLacy IH. Connections among explanatory strategies used to examine genotypic variety and genotype-by-condition collaboration in plant rearing multi-condition tests. Theor. Appl. Genet. 1994; 88:561-572. DOI: 10.1007/BF01240919
11. Crossa J, Cornelius PL. Locales relapse and moved multiplicative model grouping of cultivar preliminary destinations under heterogeneity of blunder fluctuations. Harvest Sci. 1997; 37:406-415. DOI: 10.2135/cropsci1997.0011183X003700020017x
12. Crossa J, Cornelius PL, Sayre K, Ortiz-Monasterio R, Iván J. A moved multiplicative model combination technique for gathering conditions without cultivar rank change. Yield Sci. 1995; 35:54-62. DOI: 10.2135/cropsci1995.0011183X003500010010x
13. Crossa J, Yang RC, Cornelius PL. Examining hybrid genotype × condition cooperation utilizing direct bilinear models and blended models. J. Agric. Biol. Environ. Detail. 2004; 9:362-380. DOI: 10.1198/108571104X4423
14. DeLacy IH, Basford KE, Cooper M, Bull JK, McLaren CG. Investigation of multi-condition preliminaries, a chronicled viewpoint," in Plant Adaptation and Crop Improvement, eds M. Cooper and G. L. Sledge (Wallingford: CABI), 1996, 39124.
15. Drummond ST, Sudduth KA, Joshi A, Birrell SJ, Kitchen NR. Factual and neural techniques for site explicit yield forecast. Trans. ASAE. 2003; 46:5. DOI: 10.13031/2013.12541
16. Burgueño J, Crossa J, Cotes JM, Vicente FS, Das B. Expectation evaluation of direct blended models for multienvironment preliminaries. Harvest Sci. 2011; 51:944-954. DOI: 10.2135/cropsci2010.07.0403