

# Knn based Crop and Fertilizer Prediction



Prakash U.M, Sristika Bora, Abhishek Gautam, K.R. Gokul Anand

**Abstract:** India has always been active in agriculture, in fact even in this age of industrialization agriculture and agriculture-based industries continue to be a main source of income for a large percentage of the population. Machine learning and data mining have become, in the present day, are very important mediums when it comes to research in the crop yielding domain. Many a times we come across news on the paper about farmers committing suicide because of crop failures and increase in loans. In preventing such situations, crop yield prediction software can play a very important role. This research is an attempt in proposing a method to predict the success of crop for a particular area by using data on amounts and ratios of different components of soil like nitrogen, potassium, phosphorus and environmental statistics on temperature and weather. Various machine learning algorithms are used to get an accurate result. KNN is used for classification and regression prediction problem. It also attempts in providing a precise output on what fertilizers can be used to better the yield. Through this, therefore, farmers will also be able to predict their profits and final revenues.

**Keywords:** KNN, Web Designing, JavaScript, HTML.

## I. INTRODUCTION

India being such a highly populated country with random changes of weather conditions is in a great need for protection of the food resources. Failure in crops affect the population drastically. Different kinds of crops need different kinds of conditions. For example, sugarcane need hot tropical areas with enough rainfall while oranges need winter season. Also, there are as many as 15 different agro-climatic regions in our country with each region specialized in different kinds of crops. Farmers hence need an efficient way to gather information about what kind of crops will be suitable for their land and also to be prepared for adverse conditions. A proper software which gives a precise output can make a huge difference. By using data mining, image processing and ML advances can be made in the crop yielding prediction domain. This is an attempt in making a software where inputting data on the components of the soil, ph., and environmental factors

## II. RELATED WORK

India has always been a country where agroindustry's are very important source of living for the people. In India many farmers face financial loss due to natural calamities.

Prediction of the crop yield before time can be considered a very important way which can help farmers and government. Predicting a crop well in advance require different data sets sodium, phosphorus, potassium, ph level, temperature, etc. This paper goes through the data mining techniques that can be used to predict the yield of crop. Gives a detailed description on various algorithms used. Also discusses accuracy and recommendation. World population is increasing day by day and we need a proper crop security. This paper discusses on the various deep learning techniques that can be used to predict the yield of crops. It is based on a worldwide view. Since world population is continuously growing crop security is becoming more and more important. Monitoring the yield is therefore very essential. This study focuses on soybean yield in Lauderdale, USA. Focuses on using 3d CNN model using spatiotemporal characteristics. RMS is used to compare the results with other such techniques. Agriculture plays a very important role in developing the GDP of India. Crops can be very responsive to change in temperature or weather conditions. In this paper they have discussed about some method that can be used for the forecasting. These methods are fuzzy based, namely the auto regressive moving average, ARMAX which is the same as ARMA but with exonerated variables and SARIMA which goes by seasonal auto regressive integrated moving average. All the three are compared and the best of all is chosen. like temperature, weather etc. will result in the output of the 1. [4] Researchers believe that the population of the world is kind of crops that are best suitable for the conditions and the kind of fertilizers that can be used to maximize the output. Basically, the main goal is to increase the profitability and the sustainability of the crops. So, farmers will no longer have to predict the yield simply on the basis of their previous going to get doubled in around 50 years of time. In this paper A method of prediction has been proposed by the researcher which uses the k means algorithm and also KNN. Classification and clustering are done using the MATLAB and the WEKA. A comparison has been made using various different features and according to that the result has been proposed that this method will generate better results as compared to the data mining traditional ways of prediction of yield. In this paper for the prediction of crop, authors use the Random forests (RF). They also use the multiple linear regression (MLR) for the best results. They used data from various crop yield for training and testing. RF was found very good in predicting crop yield. In this paper, we mainly focus on the prediction of rice yield using the different techniques of data mining. Rice yield nearly contribute to 40% of the total yield generation. This paper uses different types of data mining techniques for the Maharashtra state government.

Revised Manuscript Received on April 18, 2020.

\* Correspondence Author

**Sristika Bora\***, Dept. of CSE, SRM Institute of Science and Technology, Chennai, India, Email: borasristika@gmail.com

**Prakash U.M.**, Assistant Professor, Dept. of CSE, SRM Institute of Science and Technology, Chennai, India, Email: prakashm3@srmist.edu.in

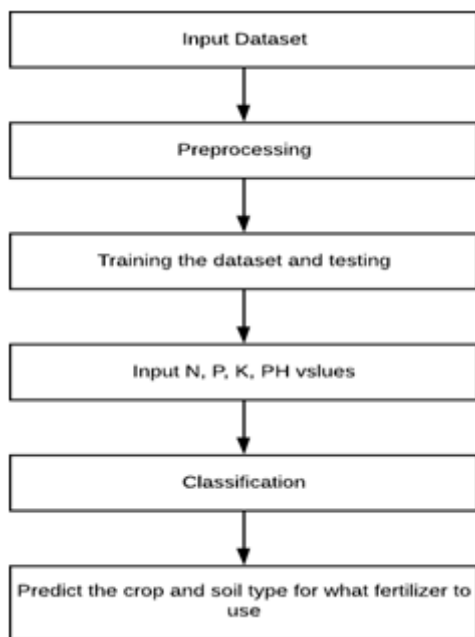
**Abhishek Gautam**, Dept. of CSE, SRM Institute of Science and Technology, Chennai, India, Email: sagargautam85@gmail.com

**K.R. Gokul Anand**, Assistant Professor, Dept. of CSE, SRM Institute of Science and Technology, Chennai, India

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Developing the better strategies for the yield prediction can help the farmers as well as the government agricultural funds. In India, agriculture almost gives the 50% employment to the people of the country. The agriculture also contributes to 18% of the total GDP of the country. But in spite of all of this the farmers are in great debt and doing suicide because of the weather unpredictability and loss in the agriculture. Soil type and different factors are directly influenced the crop yield production. This paper focuses on the crop yield by using different agriculture dataset with the different data mining techniques.

## III. IMPLEMENTATION



**Fig I Architecture Diagram**

Access to quality and mechanized input/outputs is a very big challenge of existing system and largely affect the yielding of crops and also the profitability. In the proposed system we are going to make the user interface easier to understand for farmers. We not only recommend one single crop like the previous systems but also provide 2-3 alternatives so that the farmer can choose according to his needs or financial condition. We also attempt to give an idea on how much profit will be attained and if there are any danger of adverse situations and how to confront them. A detailed description is given about that particular crop, fertilizer to use and also the types that exist. Many of the earlier systems were only based on data mining algorithms but here we will use machine

### A. Data Preprocessing

This process is basically where the raw data is converted to a form that is properly understandable. Since real world data is often too messy, incomplete or inconsistent, the raw data is cleaned and metadata is appended to it so that the data is easy to train. The metadata is loaded and added to the data. There are three parts, data cleaning, data transformation and data reduction. The following data is divided into two categories i.e training and testing data. In order to split the data, we import `train_test_split` which will split according to the

given weight that is required for the algorithm. The percent of train and test is 20 and 80 respectively.

### B. Model Creation

We create data into two models:

Training model Testing model

The ratio of the test and training set is 1:4. That is, if there are 20 of training pictures, there will be 80 of test pictures. techniques used in this process. Result for the same is compared in fig3.

### C. Model Evaluation:

We apply the machine learning algorithm for testing part and get the accuracy of this model.

### D. Prediction:

Then we have the graphical user interface which is made by using bootstrap. The page asks us about the values of phosphorus, nitrogen, potassium, PH, temperature, season etc. The data that will be entered will be compared to the dataset values, so that the crop and fertilizer details can be outputted.

**E. Methodology:** We use a map to take in the temperature value. The soil properties are inputted manually.

**F. Dataset description:** The data that is stored in the first table if basically used to predict what kind of crops that we have taken into our work. We have started with 8 crops. These crops include peas, rice, garlic, sugarcane, orange, potato, onion and tomato. Then comes the table which we use for the training, this table has various entries of each of the crops at different conditions. The more varied, the better precision. We use another table which has the fertilizers and is used to predict the perfect fertilizer to use for the resultant crop. The user will input data on the mounts of nitrogen, potassium, Sulphur, ph. level, temperature and whether conditions.

Packages used: numpy, Matplotlib pyplot, TensorFlow, Scikit-learn, Jupiter.

### G. Algorithm:

Before choosing a algorithm we looked at following factors- ease to interpret input, calculation time, prediction power.

We use spatial spectral classification and AdaBoost classifier. Also, KNN which is given below.

KNN Pseudo Code:

We can implement a KNN model by using the below steps:

1. First of all we load the data from source.
2. The value of k is inserted/initialized.
3. 1 to total no. of training od the data points is then performed.

1. We use the Euclidian distance method to calculate distance. We now calculate the distance between each of the rows and the test data.
2. We then sort the distances that we calculated in ascending order.
3. We choose the topmost k rows.
4. We then choose the most frequently occurred class.
5. This is the predicted class. This is the output after the KNN.

As we see the points that are closest on the BS are all marked in the colour red. The points that are closest to the point BS can be assumed to belong to the class of

RC. In this case our choice is pretty clear because of the votes of the neighbours all going to the RC

#### IV.RESULT AND DISCUSSION

The output is a web page which shows which crop is suitable for the conditions specified. At first it starts with the introduction about the crop, then comes the required climate and soil type for the crop. It also specifies the different varieties of the particular crop type that can be grown there, and the way of propagation. Then comes a detailed description of the fertilizers and the kind of irrigation required. Finally, it ends with the way it should be harvested.



Fig V: Final Output

For understanding the KNN algorithm. In fig. III it is shown how we decide the class of the test member. We use the k nearest neighbor method which we mentioned above:

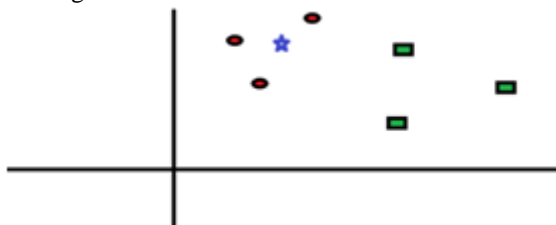


Fig III: KNN diag 1.

Now what we wish is to find where the blue star belongs. Precisely in which class. It can belong to the class of the red rounds or among the green colored squares. The letter K will be denoting the number of adjacent neighbors from whom we will be seeking votes from. In case k is 2 we take from 2 neighbors, in case of k being 3 we take from 3 neighbours. We circle the neighbours from whom we take votes from.

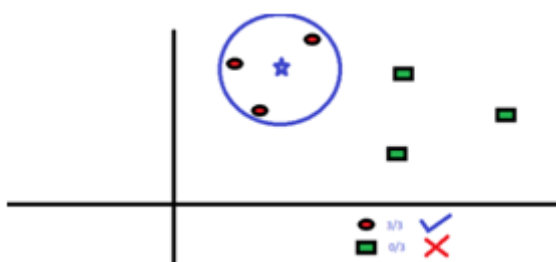


Fig IV: KNN diag 2.

#### V. CONCLUSION AND FUTURE WORK

A very precise and accurate output is given out in this process. Also, our user interface is easy to use and understand. This is a very good way a farmer can predict and plan for his upcoming plantation. We have used only 8 crops in our system. The system can be made better by using more types of crops, even trees. Many times, companies plant trees on environment day and then the trees end up withering away. Before such plantations, this system can be consulted with so that the appropriate plan type can be chosen for the particular type of soil or place. Even improvement can be done on the number of factors used in the prediction system. The webpage can contain translation option since some farmers may be more comfortable in using native language.

#### REFERENCES

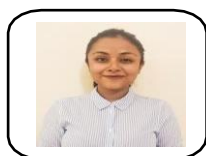
1. Yogesh gandge, sandhya " A study on various data mining techniques for crop yield prediction",IEEE International Conference,2016.
2. Anil Suat Terliksiz; d. Turgay Altılyar " Use Of Deep Neural Networks For Crop Yield Prediction: A Case Study Of Soybean Yield in Lauderdale County, Alabama, USA ", IEEE, 2018.
3. Shivam Bang, Rajat Bishnoi, Ankit Singh Chauhan, Akshay Kumar Dixit, Indu Chawla "Fuzzy Logic based Crop Yield Prediction using Temperature and Rainfall parameters predicted through ARMA, SARIMA, and ARMAX models" IEEE Xplore,2019
4. A Suresh, P. Ganesh Kumar, M. Ramalatha- Prediction of major crop yields of Tamilnadu using K-means and Modified KNN" IEEE Xplore,2018.

5. Jig Han Jeong et al , "Random Forests for Global and Regional Crop Yield Predictions ",PLOS-ONE  
DOI:10.1371/journal.pone.0156571 June 3, 2016.
6. Dakshayini Patil et al,"Rice Crop Yield Prediction using Data Mining Techniques:An Overview", International Journal of Advanced Research in Computer Science and Software Engineering ,Volume 7, Issue 5, May 2017.
7. Shruti Mishra, Priyanka Paygude, Snehal Chaudhary, Sonali Idate "Use of Data Mining in Crop Yield Prediction" IEEE Xplore Compliant - Part Number:CFP18J06-ART, ISBN:978-1- 5386-0807-4; DVD PartNumber:CFP18J06DVD , ISBN:978-1-5386-0806-7
8. .Snehal S.Dahikar, Dr.Sandeep V.Rode , "Agricultural Crop Yield Prediction Using Artificial Neural Network Approach ", International journal of innovative and research in electrical, instrumentation and control engineering, volume 2,Issue 2,2014.
9. D Ramesh, B Vishnu Vardhan, "Analysis of crop yield prediction using data mining techniques", International Journal of research in engineering and technology,eISSN:2319-1163.
10. Ramesh A. Medar, "A survey on data mining techniques for crop yield prediction", International Journal of advance in computer science and management studies, ISSN:2231-7782, volume 2, Issue 9, 2014.

## AUTHOR PROFILE



**Prakash U.M.**, Associate professor in the Department of Computer Science and Engineering at SRM Institute of Science and Technology (SRM IST). He has 7 years of teaching and research experience and published more than 12 research papers. His research interest's area includes Big data, machine learning, computer networks



**Sristika Bora**, currently pursuing Bachelor of Technology in Computer Science and Engineering at SRM Institute of Science and Technology. This paper is one of her beginning works in research area. Her main area of interest includes Machine learning, image processing, big data analysis and internet of things.



**Abhishek Gautam**, currently pursuing Bachelor of Technology in Computer Science and Engineering at SRM Institute of Science and Technology. This paper is one of his beginning works in research area. His main area of interest includes Machine learning, image processing, big data analysis and internet of things.



**K.R. Gokul Anand**, Associate Professor in the Department of Computer Science and Engineering at SRM Institute of Science and Technology (SRM IST). He has 7 years of teaching and research experience. His research interest's area includes machine learning.