

Crop Yield Prediction and Fertilizer Recommendation

¹Varshini Naresh, ²Mrs.Vatsala.B.R, ³Dr.C.Vidya Raj

¹U.G. Student, ²Assistant Professor, ³Professor, Department of Computer Science and Engineering,
The National Institute of Engineering, Mysuru, Karnataka, India, ¹varshijun7@gmail.com,

²vatsalabr@nie.ac.in, ³vidya_rajc@yahoo.co.in

Abstract—Agriculture is a major source of the economy of the country. The Prediction of crop yield helps in managing the storage of crops as well as it directs the transportation decisions, and risk management issues related to crops. Food is perhaps the most important aspect of life, without it humankind can't survive. Hence, storing food becomes an absolute compulsion for countries all over the world, especially the developing ones. The important aspect of the government is to fulfill the need of food for the long term, particularly at a time of natural disaster. There are many low-lying regions hence the estimation of food is very demanding. In this paper, we attempt to provide a precise and accurate decision in predicting crop yield and deliver the end-user with proper recommendations about the required fertilizer ratio based on parameters such as atmospheric and soil parameters of the land which enhance to increase the crop yield and increases farmer revenue.

Keywords— *Fertilizer Recommendation system, Crop yield prediction, parameters, data mining, agriculture.*

I. INTRODUCTION

India is an agriculture-based country where most of the people derive their living from this sector. Agriculture is having a great impact on the country's economy. The prediction of crop yield in advance can help the farmers and the Government bodies to plan for storage, selling, fixing minimum support price, importing /exporting etc.

Till now only the past experience of the farmer was used which consist of randomly counting the number of seed buds that a plant is having and within each seed bud the number of seed it holds. Then based on experience the farmer used to predict the crop yield. Information technology can be used to avert the risk associated with the agriculture and it can also be used to predict the crop yield more accurately prior to harvest. Yield prediction needs different kinds of data gathered from different sources like meteorological data, Agri-meteorological, soil (pH,N,P,K) data, remotely sensed data, agricultural statistics etc. To handle such a huge data the best option we have is Data Science.

Data Science is a method by which one can extract the knowledge from the huge bulk of data. We predict crop yield and also suitable required fertilizers recommended to improvise the crop yield. We use all different agriculture parameters for prediction of crop yield and fertilizers recommendation.

The system comes with a model to be precise and accurate in predicting crop yield and deliver the end user with proper recommendations about required fertilizer ratio based on

atmospheric and soil parameters of the land which enhance to increase the crop yield and increase farmer revenue.

II. LITERATURE REVIEW

Data mining approach for prediction of crop yield and analysis of soil behavior. Prediction of Yield is popular among farmers these days, which contributes to the proper selection of crops for sowing. This makes the problem of predicting the yielding of crops an interesting challenge. Earlier yield prediction was performed by considering the farmer's experience on a particular field and crop. This work presents a system, which uses data mining techniques in order to predict the category of the analyzed soil datasets. The category, thus predicted will indicate the yielding of crops. The problem of predicting the crop yield is formalized as a classification rule, where Naive Bayes and K-Nearest Neighbor methods are used.

Shruti Mishra et al 2018 have indicated that applying the data mining techniques on historical climate and crop production data several predictions are made which increase the crop productivity. The decision support system has to be implemented for the farmers to take proper decisions about soil and crop to be cultivated. They have collected the dataset with attributes of the crop season, Area and production in hectares and analyzed with various algorithms in WEKA. They analyzed data with four methods and found their accuracy and compared with each other. The four methods used are J48, IBK, LAD tree, LWL in WEKA. They concluded that the IBK had got more accuracy when

compared to all other and that depends upon the nature type and the nature of the dataset.

Effect of temperature and rainfall on paddy yield using data mining Although crop production has been coupled to a number of parameters. The work investigated the effect of temperature and rainfall variation on paddy yield. The aim was to determine the relationship between the daily temperature and actual paddy yield; daily rainfall and actual paddy yield at Ludhiana and Patiala of Punjab District, India. In addition, the research was carried out to predict the crop paddy yield at these locations within the agricultural region, given the prevalent temperature and rainfall conditions.

Data Mining is an area in which huge amounts of data are analyzed in different dimensions and angles and further categorized and then eventually summarized in to useful information. Data Mining is the process of finding correlation or patterns among dozens of fields in large databases. The research is conducted taking under consideration the various stages of the paddy plant that are vegetative stage, reproductive stage, maturation stage and grain filling stage. System discovers the relationships between temperature, rainfall and paddy plant during all these stages using data mining technique “Association Rules Mining”. Data collected from government sector and we can make use of synthetic data for analysis and to predict the patterns between temperature, rainfall and paddy plant.

Eswari et al 2018 have indicated that yield of the crop depends on the perception, average, minimum and maximum temperature. Apart, from that, they have taken one more attribute named crop evapotranspiration. The crop evapotranspiration is a function of both the weather and growth stage of the plant. This attribute is taken into consideration to get a good decision on the yield of the groups. They all collected the dataset with these attributes and send as input to the Bayesian network and classify into the two classes named true and false classes and compared with the observed classifications in the model with a confusion matrix and bring the accuracy. Finally, they concluded that crop yield prediction with Naïve Bayes and Bayesian network give high accuracy.

III. ALGORITHM OVERVIEW

Prediction of agriculture yield is essential to deal with storage of crops, transportation decisions and risk management issues related to crops. Current system is manual where we compare the previous results with the present. Based on the previous experiences and results we come to know how much crop yield will be produced. There is no automation to predict the crop yield using the constraints temperature, rainfall, humidity, area and region.

Problems of existing system: The major problems faced in the present available system is: Manual Process, Time Consuming, Less Reliable, Less Efficient, Less User Satisfaction

A. Methodology

The proposed work can be demonstrated using the Visual Studio platform which can easily work on a normal PC or laptop with 4GBs of RAM. We use DOTNET framework and uses the programming language C#.NET .MS SQL server as backend to store database of yield prediction of previous years.

This system uses different agriculture parameters for prediction of crop yield and fertilizer recommendation. Here the results are very precise and accurate provide the end user with recommendation of fertilizers which increases the yield and also the farmers revenue. The major steps necessary for obtaining the accurate results are as follows:

- **Selection of agriculture field:** Consider any agriculture field for the crop yield Prediction system.
- **Selection of crop:** Consider any crop of choice which will be shown in that field.
- **Input data:** Data may include information regarding soil (Nitrogen(N), Phosphorus(P), Potassium(K) Content, micronutrients present in soil, moisture in soil etc) which is collected over some period of time.
- **Pre-processing:** Data which is collected should be pre-processed.
- **Attribute Selection:** Important Features have to be extracted.
- **Classification Algorithm:** Two efficient algorithm as been employed. Naive Bayes Algorithm for Crop yield prediction. This algorithm provides us high accuracy and KNN Algorithm is used for Fertilizer Recommendation.

Naïve Bayes Algorithm formulae:

$$P(\text{attribute value}(a_i)/\text{subject value}(v_j)) = \frac{(n_c + mp)}{(n+m)}$$

KNN Algorithm formulae:

$$\text{Distance} = \text{Sqrt} [(q_1-p_1)^2+(q_2-p_2)^2+(q_3-p_3)^2]$$

- **Result:** Prediction and recommendation can be provided to the farmers based on the results obtained.

B. Expected Outcomes

This system is an agriculture application which analyses the previous data related to rainfall, temperature, humidity and other factors and crop yield. Proposed system makes use of data mining in agriculture for decision making. The work is conducted taking under consideration the various constraints such as temperature, rainfall, humidity, region, area and

other constraints. System uses Naive Bayes Algorithm for crop yield prediction and KNN algorithm for suitable fertilizers recommendations. Data collected from government sector (region, temperature, rain, humidity and yield prediction based on year wise and location wise) can be used to predict crop yield and suitable fertilizers to improvise yield.

Advantages of proposed system:

The proposed system is useful for agriculture department to predict crop yield and to suggest the suitable fertilizers if yield is low. It is useful to farmers to know the crop yield and required fertilizers to improvise yield. In this proposed system there is no need to analyse manually. All records are stored on server for fast and easy access.

System uses DOTNET framework and uses the programming language C#.NET and we use MS SQL server as backend to store database of yield prediction of previous years.

Naive Bayes algorithm is a collection of classification algorithms based on Bayes' Theorem. In this algorithm we predict the probability of different classes based on various parameters.

Algorithm 1: Naïve Bayes Algorithm

1. Create an array list "s" which is used to store the output.
2. Create string parameters such as features, tab Features and store all the parameters into tab Features. Value of m and p is calculated.
3. The store all the parameter values of testing dataset into an array list using a for loop.
4. Fetch the training dataset and execute the formulae of Naïve Bayes.

$\text{double } x = m * p;$

$\text{double } y = n + m;$

$\text{double } z = nc + x;$

$\text{result} = z / y;$

5. Multiply the result by p. Than, create a temporary array list called list1.
6. Sort and Reverse. When we reverse the highest value is stored in the first index 0.
7. Since 0th Index as the highest value now we need to check this with the output. The corresponding array list in s is the output.

K-Nearest Neighbour is one of the simplest algorithms in Machine Learning, based on Supervised Learning technique.

This algorithm assumes the similarity between the new data and available records and put the new record into the category that is most similar to the available records. This algorithm stores all the previous data and classifies a new data based on the similar records. This means when new record appears then it can be easily classified into a similar category of records using K- NN algorithm.

Algorithm 2: KNN Algorithm

1. Create array list Distance, RecordID , s .The possible outcomes are fetched and is stored in s. Initiate a value to m, where m is k. K determines the nearest neighbours.
2. Distance between the parameters of record is calculated using Euclidean distance formulae.
3. Create a temporary array list and store all the distance values in the array list.
4. Sort, Identify the least two numbers and we need to find which index does the least two numbers belong.
5. Calculate the number of possible outcomes. Create a temporary array list to store the record
6. Sort and Reverse. When we reverse 0th index will have the highest number.
7. Since, 0th index will have the highest value that is the outcome of the record.

IV. CONCLUSION

The current agriculture field is facing many problems and one of the major problems that the farmers are facing is lack of profit level. Farmers grow the crops but do not get the proper yield which leads to less profit level. Yield prediction pays vital role in the agriculture department. It is important to predict the yield early, based on parameters such as temperature, rainfall, soil features etc. This project results with the precise and accurate crop yield and deliver the end user with proper recommendations about required fertilizer ratio based on atmospheric and soil parameters of the land which enhances the crop yield and increases farmers revenue.

REFERENCES

- [1] S.Bhanumathi,M.Vineeth and N.Rohit," Crop yield prediction and efficient use of fertilizers",IEEE International Conference on Communication and signal processing, April 4-6,2019.
- [2] Kuljit Kaur,Kanwalpreet Singh Attwal ," Effect of temperature and rainfall on paddy yield using data mining ",IEEE International Conference on Cloud Computing, Data Science and Engineering, 2017.

- [3] D Ramesh, B Vishnu Vardhan, "Region specific crop yield Analysis : "UACEE International Journal of Advances in computer Science and its Applications-IJCSIA volume 3:issue 2.
- [4] S.Veenadhari, Dr Bharat Mishra, Dr C D Singh, "Soyabean productivity modeling using Decision Tree Algorithms," International Journal of Computer Applications Vol 27-No7 Aug 2011.
- [5] George Rub, "Data Mining of Agricultural Yield Data:A comparison of Regression models.
- [6] Wim G.M. Bastiaanssen, Samia Ali, "A new crop yield forecasting model based on satellite measurements applied across the Indus Basin,Pakistan," Agriculture, Ecosystems and Environment 94 (2003) 321–340.
- [7] Sudhanshu Sekhar Panda, Daniel P. Ames , and Suranjan Panigrahi, "Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques" , Remote Sensing 2010, 2, 673-696; doi:10.3390/rs2030673.
- [8] Machine Learning by Tom M. Mitchell, ISBN: 0070428077
- [9] 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.
- [10] 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.
- [11] 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.
- [12] Pooja M C et al, —Implementation of Crop Yield Forecasting Using Data Mining , International Research Journal of Engineering and Technology (IRJET),2018,p-ISSN: 2395-0072
- [13] 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.
- [14] Jharna Majumdar, Sneha Naraseeyappa and Shilpa Ankalaki, Analysis of agriculture data using data mining techniques: application of big data, Springer journal,2017.
- [15] S.kanaga Subba Raju et al, Demand based crop recommender system for farmers, International Conference on Technological Innovations in ICT For Agriculture and Rural Development, 2017.