International Journal of Recent Advances in Multidisciplinary Topics
Volume 2, Issue 7, July 2021
https://www.ijramt.com | ISSN (Online): 2582-7839


# Smart Agricultural Crop Prediction Using Machine Learning

G. S. Jayadeva[1], S. Shanmugasundaram[2*], N. Sharath[3], C. Sanju[4], S. M. Sidhalinga[5]
[1]*Professor, Department of Electronics and Communications Engineering, B.M.S. Institute of Technology and Management, Bangalore, India*
[2,3,4,5]*Student, Department of Electronics and Communications Engineering, B.M.S. Institute of Technology and Management, Bangalore, India*

***Abstract*: Indian economy is mainly dependent on farmer's progress, making good profit in the agriculture field, technology plays an important role. Getting higher yields and improved quality of final product a system based on machine learning is proposed. In this system analyzing, quality of soil, rain pattern, weather and temperature, the farmers are suggested best crops and its required fertilizers as a solution by which farmers will get more profit on growing system suggested crop. This system is designed as a web application which uses big data analytics, prediction analysis and other techniques to predict the most suitable and profitable crop and its required fertilizers, predicts yield per hectare and value of crop based on current market price taking into consideration of current weather and soil conditions. Thus, farmers will benefit by using our system which will improve crop productivity and profit of farmers.**

***Keywords*: Big data analytics, crop yield, knn algorithm, Machine learning, profit, soil and weather conditions.**

## 1. Introduction

Agriculture is very important because it produces food and feed which is necessity to animals and human beings. It fulfill the basic need of billions of people. It is one of the major contributors to the country's GDP and economic growth. Hence, it is widely practiced in India. Agriculture sector requires more workforce than any other sectors, nowadays there is huge decrease in the agricultural work force. So, we need to fill that huge gap by making advances in agriculture with the help of technology. Thus, agricultural advancement results in gaining more profit by the farmers. Big data analytics techniques are used to analyze data sets of temperature, humidity, rainfall and soil which are collected from the meteorological department. This type of required analysis is performed by using specific software tools, many of them available as open source. By using this tools and techniques the system will have information, by this processed information the system will take better decisions. Thus, ensuring better results. Normally, farmers can guess the final yield by their experience of growing particular crop again and again. Farmers yield prediction accuracy is low and not cost effective. To meet the food requirements of the entire population of the country and for the export of some agricultural products to other countries, it is important to practice modern methods of farming by using technology instead of practicing traditional farming methods. Modern methods allow the farmers to cultivate the crops in small area with minimum amount of water, fertilizers and pesticides, which finally produces good yield and profit to the farmers. Application which is an interface between farmer and the system. By the provided data and analyzes the final values and predicts the best crop, list of fertilizers to be used, yield per hectare and total value of cultivated crop based on the current market price to the data.

## 2. Problem Identification

The cropping pattern which shows the proportion of the area under different crops at a definite point of time is an important indicator of development and diversification of the sector. Food crops and non-food or cash crops arc the two types of crops produced by the agricultural sector of the country. As the prices of the cash crops are becoming more and more attractive therefore, more and more land have been diverted from the production of food crops into cash or commercial crops. This has been creating the problem of food crisis in the country. Thus after 50 years planning the country has failed to evolve a balanced cropping pattern leading to faulty agricultural planning and its poor implementation. Achieving maximum yield rate of crop using limited land resource is a goal of agricultural planning in an agro-based country. Antecedent determination of problems associated with crop yield indicators can help to increase yield rate of crops. Crop selector could be applicable for minimize losses when unfavorable conditions may occur and this selector could be used to maximize crop yield rate when potential exists for favorable growing conditions. Maximizing production rate of crop is an interesting research field to agro-meteorologists which play a significant role in national economic. There are two types of factors which influence yield rate of crop: first is seeds quality which can be improved by genetic development using hybridization

*Corresponding author: lucifersundar@gmail.com

technology, and second is crop selection management based on favorable or unfavorable conditions.

## 3. Existing System Vs Proposed System

*1) Existing system*

The existing systems only predict yield based on input of total area of farm land in hectares but this systems doesn't predict best suitable and profitable crop, doesn't suggest fertilizers and doesn't calculate the total value of the crop. • And also the accuracy of the existing system is low compared to our system.

*2) Proposed system*

The solution proposed allows us to predict best suitable and profitable crop, its required fertilizers, estimates yield, and calculates the total value of the crop.

## 4. System Architecture

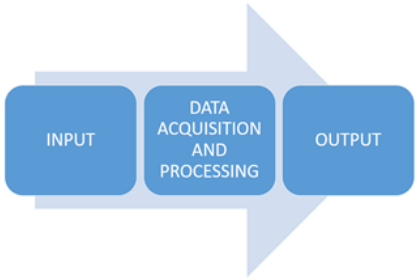


Fig. 1. Basic project flow

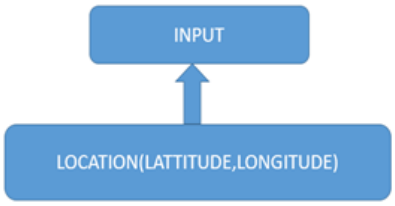


Fig. 2. Input Requirement

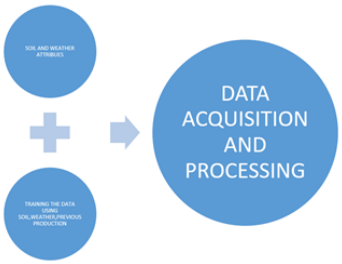


Fig. 3. Data Acquisition

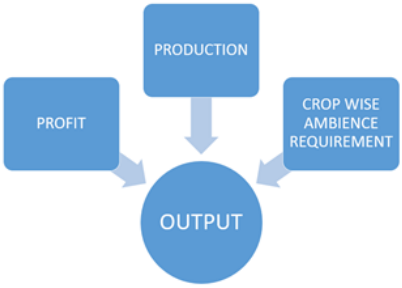


Fig. 4. Output (System Architecture)

*1) Datasets*

The data set is taken from UCI Machine Learning Repository for Smart Agricultural Crop Prediction Using Machine Learning It contains a number locations such as states,. Only physicochemical variables and the output variable districts and villages are available and contain more information about different types of crops, contains PH value of soil, weather conditions, longitude and latitude, selling prices etc. because of privacy and logistical problems.

*2) Data processing*

Data preprocessing is a method to explore data that involves raw data to be converted into accessible format. Data preprocessing is a established technique to overcome any data set issues. The preprocessing controls for missing values, normalizes the numerical data to minimize the number of variables, convert numerical data to categorical values where appropriate, etc. If all the attributes in the dataset are numerical, there is no need of any encoding. This work also check for the presence of duplicates on the dataset and take action accordingly.

The following are the specific steps in pre-processing data:

- *Cleaning of data:* data may have multiple sections that are insignificant and many missing. Data are cleaned to handle this portion. It requires the handling of missing data, noisy data, etc. Noisy data is not computer interpretable data. Due to errors in data collection, entry errors, etc.it can be created..
- *Data integration:* Data integration is a processing process which combines data from a variety of sources and allows the user to see the data uniformly.
- *Data transformation:* this is a step towards modifying data for mining in correct ways. It requires standardization, attributes selection, flexibility etc.
- *Data Reduction:* Data reduction is carried out to reduce the size of the data collection by taking only data features applicable to the mission into consideration. Data reduction technology aims to improve storage capacity, raising data storage and computational costs by operating with a vast volume of data and a wide variety of features. It comprises reductions in size, aggregation of the data cube, numbers, reduction, etc.].
- *Data Discretizing:* The division of values into bins, i.e. the decrease of the number of possible values, is the method of discretizing data. Data discretization the buckets are known to be organized and discrete values. Columns of numerical and strings can be debunked.

## 5. Machine Learning Algorithm

*1) KNN algorithm*

K-nearest neighbors (KNN) algorithm is a type of supervised ML algorithm which can be used for both classification as well as regression predictive problems. However, it is mainly used for classification predictive problems in industry. The following two properties would define KNN well −

- *Lazy learning algorithm* − KNN is a lazy learning

algorithm because it does not have a specialized training phase and uses all the data for training while classification.

- *Non-parametric learning algorithm* − KNN is also a non-parametric learning algorithm because it doesn't assume anything about the underlying data.

K-nearest neighbors (KNN) algorithm uses 'feature similarity' to predict the values of new data points which further means that the new data point will be assigned a value based on how closely it matches the points in the training set. We can understand its working with the help of following steps −

- Step 1 − for implementing any algorithm, we need dataset. So during the first step of KNN, we must load the training as well as test data.
- Step 2 − Next, we need to choose the value of K i.e. the nearest data points. K can be any integer.
- Step 3 − For each point in the test data do the following
- 3.1 − Calculate the distance between test data and each row of training data with the help of any of the method namely: Euclidean, Manhattan or Hamming distance. The most commonly used method to calculate distance is Euclidean.
- 3.2 − Now, based on the distance value, sort them in ascending order.
- 3.3 − Next, it will choose the top K rows from the sorted array.
- 3.4 − Now, it will assign a class to the test point based on most frequent class of these rows.
- Step 4 − End

## 6. Implementation

- *Step 1:* Getting all the basic libraries that are useful for the project.
- *Step 2*: Searching and getting the data set from the online sites here data.gov.
- *Step 3*: Importing the data set to the project directory.
- *Step 4:* Purify the data
- *Step 5:* Splitting the data set and applying the machine learning.
- *Step 6:* Getting the result in the application created by using the python

Steps in Execution

- Import the libraries.
- Import the data set.
- Get the basic information about the data set.
- Get the correlation parameter for the variables,
- If required plot and analyses the parameters.
- And also need to purify the data set.
- Check for nulls in the data set.
- If any drop those rows.
- Check for any special characters in the columns.
- Just convert them in to nulls and drop those rows.
- New need to select the features for the project.
- Taking all the column as input and production as output.
- Now need to split the data set in to train and test data SET. 70 in to train data and 30 test data set.
- And train the machine using the KNN model.
- And finally using this trained model we can create an application to work.
- To take the inputs from the user and display the result

## 7. Result

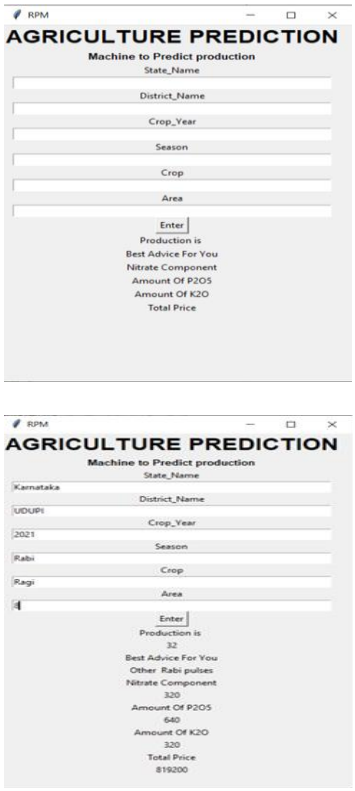


Fig. 2. Screen shots of input and output

The above result shows before and after the inputs given by the user. The first screenshot of the interface asks the user to give the details Such as state name, district name, crop year, season, the crop name and the area of the land. The second screenshot gives the results in accordance to the inputs given by the user, different results shown are: production in quintals best crop according to the conditions specified, the nitrate component, amount of phosphorous and potassium and the total price of production.

## 8. Conclusion

This paper presented an overview of Smart agricultural crop prediction using machine learning

## References

[1] DhivyaB ,Manjula , BharathMadhumathi,"A Survey on Crop Yield Prediction based on Agricultural Data", International Conferencence in Modern Science and Engineering ,March 2017.

[2] Crop Variety Selection Method using machine Learning RMK engineering college thrivallur Tamil Nadu.

[3] Smart farming using machine learning dept of Cse VJTI college, Mumbai, Maharastra

[4] Machine Learning Applications for Precision Agriculture: A Comprehensive Review, university of petroleum and energy studies, Dehradun, Uttarakhand

[5] Analysis of Crop Yield Prediction by using Machine Learning Algorithms CSE Dept, CCET, Valanchery, Kuttipura