

# A Review on Forecasting Agricultural Demand and Supply with Crop Price Estimation Using Machine Learning Methodologies

Pallavi Shankarrao Mahore, Dr.Aashish A. Bardekar

Sipna College of Engineering and Technology, Amravati, Maharashtra, India

## ABSTRACT

### Article Info

Volume 7, Issue 3

Page Number: 570-575

### Publication Issue :

May-June-2021

### Article History

Accepted : 18 June 2021

Published : 25 June 2021

Agriculture plays a vital role in Indian economy. It contributes 18% of total India's GDP. In India, most of the crops are solely dependent upon weather conditions. Hence, more yield of crops can be achieved by analyzing agro-climate data using machine learning techniques. Machine learning (ML) is a crucial perspective for acquiring real-world and operative solution for crop yield issue. From a given set of predictors, ML can predict a target/outcome by using Supervised Learning. To get the desired outputs need to generate a suitable function by set of some variables which will map the input variable to the aim output. Crop yield prediction incorporates forecasting the yield of the crop from past historical data which includes factors such as temperature, humidity, ph, rainfall, crop name. It gives us an idea for the finest predicted crop which will be cultivate in the field weather conditions. These predictions can be done by a machine learning algorithm called Random Forest. It will attain the crop prediction with best accurate value. The algorithm random forest is used to give the best crop yield model by considering least number of models. It is very useful to predict the yield of the crop in agriculture sector.

**Keywords:** Supervised Learning, Random Forest, K Nearest Neighbor, crop yield prediction, machine learning classifiers.

## I. INTRODUCTION

Agriculture is the main occupation in India and economy of the country is entirely depended on it for rural based existence. Because of certain components like atmosphere changes, unpredicted precipitation, decline of water level, utilization of pesticides unnecessarily and so on. The degree of farming in India is diminished. The accuracy of 87% is achieved

from the system and high correlation is seen between yield of crop and the climate. Agriculture inputs like chemicals, pest, soil quality and many more inputs were not considered for change in agriculture from field to field. This model is going to help farmers to make better decisions as to decide which crop to plant. Based on the seasons climate it will help farmers to make important decisions, such as import,

export, pricing, marketing before the crop is harvested.

As the time passes the need for production has been increasingly exponentially. In order to produce in mass quantity people are using technology in a wrong way. New kinds of hybrid varieties are produced day by day. However, these varieties do not provide the essential contents as naturally produced crop. These unnatural techniques spoil the soil. It all leads to further environmental harm. Most of these Unnatural techniques are used to avoid losses. But when the producers of these crops know the accurate information on the crop yield it minimizes the loss. To achieve this project is made. Using past information on weather, temperature and several other factors the information is given [1].

To overcome this challenge, an end-to-end supply chain system is presented factored in with the help of machine learning algorithms which not only promises a fair price to farmers for their produce before sowing but also provides supply of yield insights to the smart recommendation for the consumers. Predicting a crop well in advance requires a systematic study of huge data coming from various variables like temperature, humidity, rainfall, crop name soil quality etc.

Machine learning is one of the best strategies for predicting unknown values. This work manages the expectation of crop yield using machine learning. Today Machine learning plays an important role in solving various classification, forecasting and analysis problems in various fields, like in analyzing soil type and further, it can be useful for crop prediction based on the previous crop sequences in the same farmland with the current soil nutrient information. SVM (Support vector machine) plays an important role in classification as it provides best partition among classes. Predicting a crop in advance necessitates a systematic study of vast amounts of data derived from various variables such as temperature, humidity, rainfall, crop name, soil quality, and so on. Because

crop prediction involves a large number of databases, this prediction system is an ideal candidate for the application of a machine learning model. Many classification methods are also applied to get the maximum yield of crops. [2]

Data mining is machine learning tool which is used to view data in all possible ways and analyze. After analyzing the data, it is used to predict for future purposes. It can be used in several fields. These patterns provide information about crop. The aim is to increase the yield and profit for producers. The proposed system concentrates on yield, weather predictions and crop type. The dataset is taken on agriculture statistics. This dataset is used as an experimental basis. After the data processing it is divided into training and testing.

## II. LITERATURE SURVEY

**Apipattanavis, Podest'a, and Rajagopalan** [1] says Numerous research has been performed in this area, which includes weather and soil data for suitable crop identification and crop yield prediction. A stochastic seasonal weather forecasting model is presented by Apipattanavis et al., which is integrated with crop simulation model to predict expected yield and economic return. The model can also help to identify production risk associated with dry and wet climate. The implementation of this model is done for Maize crop in two regions of Argentina. An ensemble seasonal forecasting model is implemented in Europe to predict crop yield and to plan crop management activities beforehand. The model is employed for wheat crop and shows better prediction accuracy than given by the operational system. It also uses probability distribution function to shows the reliability of prediction for a particular region.

**Balamurugan et al.** [2] have implemented crop yield prediction using only random forest classifier. Various features like rainfall, temperature and season were taken into account to predict the crop yield.

Other machine learning algorithms were not applied to the datasets. With the absence of other algorithms, comparison and quantification were missing thus unable to provide the apt algorithm.

**Mishra et al.** [3], has theoretically described various machine learning techniques that can be applied in various forecasting areas. However, their work fails to implement any algorithms and thus cannot provide a clear insight into the practicality of the proposed work.

**Manjulas et al.**[4] research aimed to propose and implement a rule-based system to predict the crop yield production from the collection of past data by applying association rule mining on agriculture data from 2000 to 2012. The dataset used in this research is limited to the southern district of India thus limiting its scope for pan India implementation.

**Dahikar and Rode**[5] in their research provided the datasets of different features and applied Artificial Neural Networks for crop yield prediction. However, Dahikar fails to give a practical implementation of his proposed work. Further described how the old farming data can be utilized to depict the future expectation of harvests and yield. It likewise proposes the ranchers about what kind of yield can be developed utilizing the climate station data and gives the appropriate data to incline toward the precise season for greatness cultivating.

**Bendre, & hool** [6] describes the forecast of yield utilizing IOT with the reasonable climatic conditions and the potential outcomes of progress and its application. They have utilized the Hadoop record framework. To manufacture an expectation framework for crops and to distinguish the nuisances the characterization, investigation and forecast calculation is utilized.

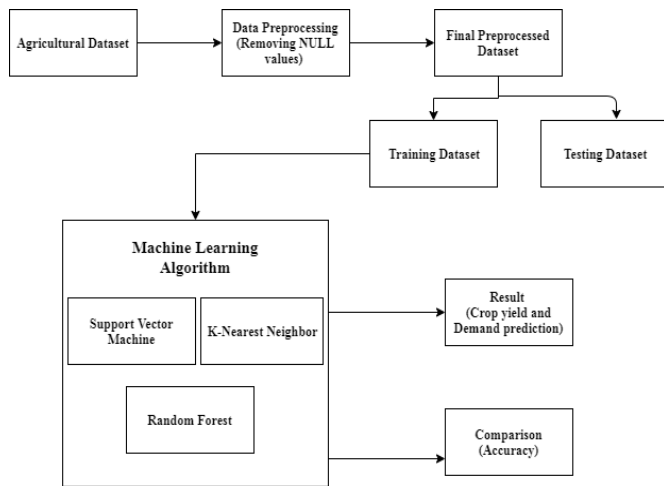
**Monali Paul, Santosh K. Vishwakarma & Ashok Verma** [7] describes the different components that associated with ecological parameters which impact the yield of the harvest are Area under Cultivation, Annual Rainfall, and Food Price Index and set up the

relationship among these parameters. The curse on the harvest yield is broke down by utilizing different ecological elements and Regression Analysis (RA), Linear Regression (LR) Algorithm.

**Dr. Rajini Kanth & Jeevan Nagendra Kumar**[8], describe the diverse mining procedures to consider crops that are quantitative and relationship them for interseason development. Grouping enormous information is a test, so k implies calculation is utilized to oversee huge information. Proper calculation is utilized to decide the harvests are chosen as incessant thing set. What's more, they centre around the administration strategies and the trimming practices of edges. a crop yield prediction model using support vector regression is presented. First, the model predicts rainfall using modular artificial neural network and then based upon amount of rainfall and some other features, the crop yield prediction is done for rice, maize, ragi and bajra crop in Visakhapatnam. A soil classification method is presented by Paul et al. , which classifies soil into three categories: low, medium and high based upon its nutrient content . Soil parameters such as pH, organic carbon, nitrogen, phosphorus, potassium and some other micro-nutrients are used for classification using KNN. The class of soil helps in identifying suitable crops for soil that gives maximum yield.

### III. PROPOSED METHODOLOGY

Machine learning is an important decision support tool for crop yield prediction, including supporting decisions on what crops to grow and what to do during the growing season of the crops. Several machine learning algorithms have been applied to support crop yield prediction research.



**Fig. 1 : Block Diagram Of Crop Yield And Demand Prediction**

Above architecture clearly explains about how the components of the system communicate among themselves starting from preprocessing of data. This proposed framework is able to finding out the crop yield. This model gives clear picture of huge amount of data capture and preprocessing of data to remove the unwanted data such as NULL etc. presented in it. During preprocessing step we split the dataset into training and testing dataset. Train dataset to detect the crop yield present in the dataset using appropriate supervised learning algorithms. Apply the machine learning techniques which are helpful for finding crop yield for any of new data occurred in the data. After this data acquisition suitable machine learning algorithm must be applied to compute efficiency and capability of the model, here we have applied various machine learning algorithms like support vector machine, K- nearest neighbor, random forest etc. Metrics like accuracy, precision will be calculated for the proposed model.

Dataset consists of few important attributes such as temperature, rainfall, humidity, ph., applied random forest algorithm for classification and regression tasks. Tried to train the model with decision trees but found that random forest algorithm reduced the over fitting problem and also improved accuracy and used

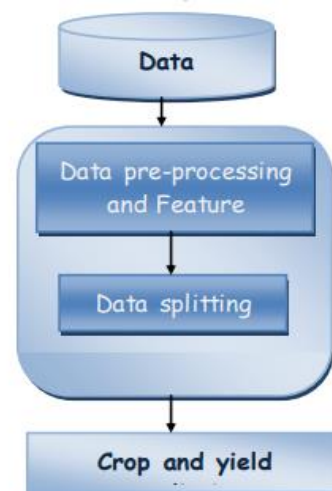
SVR, Random forest and random forest got more accuracy. The dataset which used is imported from online repository.

#### IV. PROPOSED SYSTEM

The proposed system follows Naive Bayes classifier, the supervised learning algorithm consist of the four levels to calculated and predict the crop for the suitable climate in phenomenon such as,

**A. Data Collection:** Data is composed from a different source and optimized for data sets. And the data is used to evaluate descriptively. Several abstract online outlets, like Kaggle, Google weather forestation and data government, provide the data for up to 10years in series. The data sets such as soil nature, climatic conditions and seed data are used for the crop prediction and better crop yields.

**B. Preprocessing Step:** Preprocessing the data is considered as a significant step machine learning phase. Preprocessing involves adding the missing values, the correct set of data, and extracting the functionality. Data set form is important to the process of analysis. The data collected in this step will induced in Google Colab platform in the form of python programming in order to get the desired output.

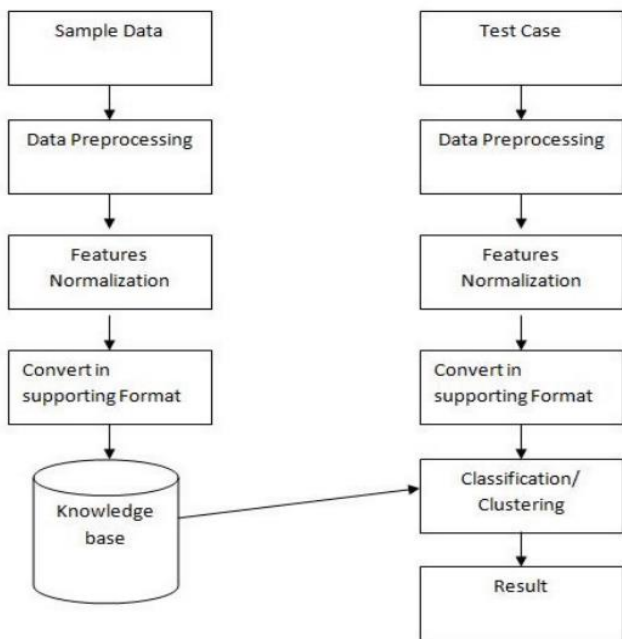


**Fig. 2 : Steps in proposed system**

**C. Feature Extraction:** Extraction of the features would reduce the data size involved to characterize a wide collection of data. The characteristics of soil, crop and weather collected from the pretreatment process establish the final training data collection. This approach selects the features based on the correlation matrix i.e. the features that has more correlation value is selected as an important predictive function for yield as shown in Fig.2

Here first we collect the data sets and process the data and we remove if there are any impurities in the data sets. Next the data is normalized if needed like it can be converted to smaller volume of data. Next the data is converted to supporting format. And then it is stored in the databases. Next the required method is applied. Now we get the final results.

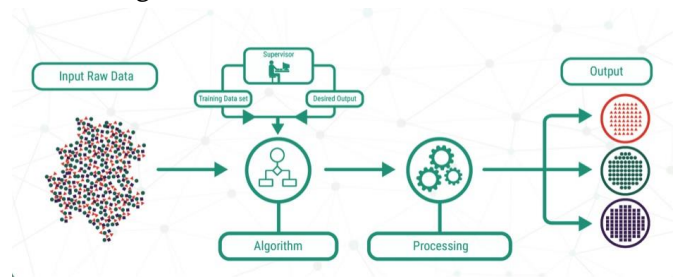
**V. WORKING METHEDODOLOGY**



**Fig. 3: System Architecture for Crop Yield Prediction**

We had a small portion of null values in the production column. Due to the miniature amount, we dropped the fields off of the dataset. This did not affect much, as the amount was minimal. We also

performed feature extraction on the dataset. Some of our data were categorical and some were continuous numeric. This type of mixed data always causes problems to the algorithms. Hence, we performed standard feature scaling on all of the data to bring them into a common scale. Feature scaling is extremely important due to the fact that, most algorithms feature a lot of internal calculation. Additionally, feature selection was done to reduce over fitting issues.



**Fig. 4: Prediction Process**

The dataset, used 80% of data for training the model and 20% of data for testing the results to obtain better results and trained the model by applying random forest algorithm. Then, compared the predicted result with the original data set. Later, estimated the accuracy of the model using test samples. Likewise, predicted the accuracy of the model with different algorithms. Then finally concluded that random forest algorithm gives us more accuracy. Hence, used random forest algorithm to train the model.

**VI.CONCLUSION**

This focuses on yield prediction according to the crop sown prediction process. Yield is being predicted for crops and crops with the help of environmental condition, along with soil condition. The various machine learning algorithms will predict the yield of the crop on the basis of temperature, rainfall, land type and soil condition etc. Training will be conducted based on dataset.



## VII. REFERENCES

- [1]. S. Apipattanavis, F. Bert, G. Podest'a, and B. Rajagopalan, "Linking weather generators and crop models for assessment of climate forecast outcomes," *Agricultural and forest meteorology*, vol. 150, no. 2, pp. 166–174, 2010.
- [2]. Priya, P., Muthaiah, U., Balamurugan, M. *International Journal of Engineering Sciences Research Technology Predicting Yield of the Crop Using Machine Learning Algorithm.*
- [3]. Mishra, S., Mishra, D., Santra, G. H. (2016). Applications of machine learning techniques in agricultural crop production: a review paper. *Indian J. Sci. Technol*,9(38), 1-14.
- [4]. Manjula, E., Djodiltachoumy, S. (2017). A Model for Prediction of Crop Yield. *International Journal of Computational Intelligence and Informatics*,6(4), 2349-6363.
- [5]. Dahikar, S. S., Rode, S. V. (2014). Agricultural crop yield prediction using artificial neural network approach. *International journal of innovative research in electrical, electronics, instrumentation and control engineering*,2(1), 683-686.
- [6]. M.R. Bendre, R.C. Thool, V.R. Thool, "Big Data in Precision agriculture", Sept, 2015 NGCT .
- [7]. Monali Paul, Santosh K. Vishwakarma, Ashok Verma, "Analysis of Soil Behavior and Prediction of Crop Yield using Data Mining approach", 2015
- [8]. Dr. T . V. Rajini Kanth, Y. Jeevan Nagendra Kumar "GISMAP Based Spatial Analysis of Rainfall Data of Andhra Pradesh and Telangana States Using R", *International Journal of Electrical and Computer Engineering (IJECE)*, Vol 7, No 1, February 2017, Scopus Indexed Journal, ISSN: 2088-8708
- [9]. M. Paul, S. K. Vishwakarma, and A. Verma, "Analysis of soil behavior and prediction of crop yield using data mining approach," in 2015 International Conference on Computational Intelligence and Communication Networks (CICN). IEEE, 2015, pp. 766–771.

### Cite this article as :

Pallavi Shankarrao Mahore, Dr.Aashish A. Bardekar, "A Review on Forecasting Agricultural Demand and Supply with Crop Price Estimation Using Machine Learning Methodologies", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, ISSN : 2456-3307, Volume 7, Issue 3, pp.570-575, May-June-2021. Available at  
doi : <https://doi.org/10.32628/CSEIT2173169>  
Journal URL : <https://ijsrcseit.com/CSEIT2173169>