

Maximize the Yielding Rate of Crops using Machine Learning Algorithm

Yogita Masare

Department of Computer Engineering
SVKM's Institute of Technology,
Dhule, India

Sneha Mahale

Department of Computer Engineering
SVKM's Institute of Technology,
Dhule, India

Manjusha Kele

Department of Computer Engineering
SVKM's Institute of Technology,
Dhule, India

Ashvin Upadhyay

Department of Computer Engineering
SVKM's Institute of Technology,
Dhule, India

Mr. Bhushan Nandwalkar

Department of Computer Engineering
SVKM's Institute of Technology,
Dhule, India

Abstract — Agriculture is the field which plays a very important role in developing our countries economy. Agriculture is the one which gave birth to advancement. India is a rural Country and its economy largely based upon crop productivity. So we can say that agriculture can be backbone of all business in our country. Selecting of every crop is very important in the farming planning. Many researchers studied guess of yield rate of crop, guess of weather, soil classification and crop classification for agriculture planning using machine learning techniques. Many changes are required in the agriculture field to upgrade changes in our Indian economy. We can improve agriculture by using machine learning system which are applied easily on farming sector. Along with all advances in the machines and technologies used in farming, useful information about different matters also plays a significant role in it. The concept of this paper is to implement the crop selection method so that this method helps in solving many agriculture problem. This improves our Indian wealth by maximizing the yield rate of crop production.

Keyword: *Crop Selection Method, Machine Learning Algorithm, Agriculture.*

1. INTRODUCTION

The main aim of agricultural planning is to achieve maximum yield rate of crops by using limited number of land resources. Many machine learning algorithms can help in raising the production of crop yield rate. Whenever there is loss in critical conditions we can apply crop selecting method and reduce the Losses. And it can be used to gain crop yield rate in favourable conditions.

This Maximizing of yield rate helps in improving countries economy. We have some other factors that influence the crop yield rate. They are fertilizer quality and crop Selection. Selection of crops depends upon two things that are favourable and critical conditions. Many researches are

carried out to improve agricultural arrangement. The aim is to get the maximum yield of crops [3].

Many categorizing Methods are also applied to get maximum yield of crops. Machine learning can be used to improve the yield rate of crops. The method of crop Selection is concern to improve crop production.

The production of crops may depend on geographical conditions of the region like river ground, hill areas or the depth areas. Weather conditions like humidity, rainfall, temperature, cloud. Soil type may be clay, sandy, saline or peaty. Soil composition can be copper, potassium, phosphate, nitrogen, manganese, iron, calcium, ph. value or carbon and different methods of harvesting. Many parameters are used for different Crops to do different predictions. These prediction models can be studied by using researches. These predictions are classified as two types. One is traditional statistic method and other is machine learning techniques. Traditional method helps in predicting single sample spaces. And machine learning methods helps in predicting multiple predictions. We need not to consider the structure of data models in traditional method where as we need to consider the structure of data models in machine learning methods.

2. LITERATURE SURVEY

In [1] **Pavan Patil, Virendra Panpatil, Prof. Shrikant Kokate** had done tremendous work for Indian farmers by creating efficient crop recommendation system. They developed system using classifier models such as Decision Tree Classifier, KNN, and Naive Bayes Classifier. The proposed system can be used to find best time of sowing, growth of plant and Plant harvesting. They used different classifier for achieving better accuracy for example: Decision tree shows less accuracy when dataset is having more variations but Naïve Bayes provides better accuracy than decision tree for such datasets. The best advantage of system

that it can easily scalable as it is/be used to test on different crops.

In [2] **Mayank Champaneri**, **Chaitanya Chandvidkar**, **Mansing Rathod**, **Darpan Chachpara**, have concluded that this paper build improvised system for crop yield using supervised machine learning algorithms & with goal to provide easy to use User Interface, increase the accuracy of crop yield prediction, analyze different climatic parameters such as cloud cover, rainfall, temperature, etc. In the proposed system they focused on MAHARASHTRA State for implantation & for data gathering they used govt. website such as www.data.gov.in. For crop yield prediction they used algorithms such as Random Forest Algorithm & for ease of use they developed web page so that it will be easy to use for all. The main advantage of proposed system is accuracy rate is more than 75 per cent in all the crops and districts selected in the study.

In [3] **Ramesh Medak**, **Vijay S. Rajpurohit**, **Shweta**, have concluded that this paper will review that various applications of machine learning in the farming sectors. And also helps in can be select proper crop select land and select season solved using these techniques. The algorithms use are Naive Bayes and K-Nearest Neighbour. The algorithms are uses accuracy of performance.

In [4] **Nishit Jain**, **Amit Kumar**, **Suhil Garud**, **Vishal Pradhan Prajakta Kulkarni**, have concluded that this paper helps in predicting crop sequences and maximizing yield rates and making benefits to the farmers. Also, Using Machine learning applications with agriculture in predicting crop diseases, studying crop emulations, different irrigation patterns. The algorithms used are artificial neural networks. The major problem with neural network is that the appropriate network which suits best for the solution is very hard to achieve and it includes trial and error. The second problem with neural network is the hardware dependency as the algorithm includes more computations backward and forward the processing needs more. Determination of proper network structure requires experience and time. The proposed system also focuses on crop selection using environmental as well as economic factors. The system also uses the economic factor that is the price of the crop which plays a major role in case if the crops with same yield but different crop price. The system also uses other method which is crop sequencing which gives a full set of crop which can be grown throughout the season. The proposed system also focuses on crop selection using environmental as well as economic factors. The system also uses the economic factor that is the price of the crop which plays a major role in case if the crops with same yield but different crop price. The system also uses other method which is crop sequencing which gives a full set of crop which can be grown throughout the season.

In [5] **E. Manjula**, **S. Djodiltachoumy**, has have concluded that this paper helps in improving the yield rate of crops by using rule based mining. The paper uses association rule mining to predict the yield of the crop. The algorithms used are k-Means Algorithm, clustering method and A priori

association rule mining. The major disadvantage is that the paper uses association rule mining for prediction of crop yield. The problem with association rule mining is that it generates too many rules in some cases and the accuracy of the prediction reduces. Also the rules tend to vary as per dataset and the results also very greatly. The proposed system mainly focuses on the issue of yield prediction of crop which plays very important role in crop selection as farmer can select crop with maximum yield. The systems uses association rule mining to find rules and crops with maximum yield. This system focuses on creation of a prediction model which may be used to future prediction of crop yield.

In [6] **Rakesh Kumar**, **M.P. Singh**, **Prabhat Kumar** and **J.P. Singh**, has have concluded that this paper helps in improving the yield rate of crops by applying classification methods and comparing the parameters. The paper explains the use of different algorithms to achieve the same. The algorithms suggested are Bayesian algorithm, K-means Algorithm, Clustering Algorithm, Support Vector Machine. The disadvantage is that there is no proper accuracy and performance mentioned in the paper as per implementation of the suggested algorithms. The paper is a survey paper and only suggests the use of the algorithms but there is no implementation evidence provided in the paper. The method applied on this paper for crop choice focuses specially on the plants which may be grown as according to season. The proposed approach resolves choice of crop (s) primarily based on prediction yield price encouraged by parameters (e.g. Climate, soil kind, water density, crop kind). It takes crop, their sowing time, plantation days and anticipated yield fee for the season as input and finds a sequence of vegetation whose production in keeping with day are maximum over season.

3. METHODOLOGY

All we know that agriculture is the backbone of the Indian economy. But, now a days the modern people don't have awareness about the cultivation of crops at the right time and at the right place. because of these cultivating techniques the seasonal climatic condition are also being changed against the fundamental assets like soil, water, air which leads to insecurity of food. So, there are some techniques and algorithms to make accurate prediction for crops and with the help of those algorithms we can predict yield also.

We focus on main four objectives:

- 1) To use ML Algorithm to predict crop yield
- 2) To provide easy to use interface
- 3) To increase the accuracy of crop yield prediction
- 4) To analyses different climatic parameters such as rainfall, temperature, and cloud cover.

Machine learning in agriculture used to improve the production and grade of the crops in the agriculture sector. In

this system there two modules Crop yield prediction and Fertilizer prediction.

A. Crop yield prediction module:

Crop Yield Prediction is the methodology to predict the yield of the crop using different parameters like temperature, rainfall, pesticides, etc. Crop yield is a measurement of the amount of agricultural production crop. There are four main factors are crop yield are soil fertility, accessibility of water, climate, and decrease crop yield by negatively affecting the associate growth factor. Yield is return measure for an investment over a set period of time, expressed as a percentage. Yield includes price increases as well as dividends paid, calculated as the net realized return divided by the principal amount invested. There are a number of factors which are likely to have provide to sustained yield gains: fertilizer application, irrigation, increased soil tillage, and improved farming practices. In this module first user has to give the input to a system like crop name, area or district name and season for crop yield prediction. After this system is going to apply ML algorithm to the input data and it will be produced the output of it like yielding rate in Poor, moderate, good for the crop.

B. Fertilizer prediction :

Plants need to be fertilized because most soil does not produce the required nutrients required for optimum growth. It looks nutrients from the soil to build those plant tissues. By fertilizing your park, you reload lost supplements and ensure that this year's plants have the food they need to flourish. Excess fertilizer alters the soil by creating too high of a salt absorption, and this a hurt beneficial soil microorganisms. Over fertilization can lead to sudden plant extension with a short root system to supply adequate water and nutrients to the plant. So, in this module first user has to give the input to a system like crop name, area or district name and season for fertilizer prediction. After this system is going to apply ML algorithm to the input data and it will be produced the output of it like this particular fertilizer is use full or not and also how much it will being good or bad god the crop.

Dataset:

	A	B	C	D	E	F	G	H
1	State	District	Crop	Season	Year	Yield		
2	Maharashtra	Ahmednagar	Bajra	Kharif	2017	Good		
3	Maharashtra	Akola	Soybean	Zaid	2017	Poor		
4	Maharashtra	Amaravati	Wheat	Rabi	2017	Moderate		
5	Maharashtra	Aurangabad	Jowar	Kharif	2017	Poor		
6	Maharashtra	Beed	Cotton	Kharif	2017	Good		
7	Maharashtra	Bhandara	Rice	Kharif	2017	Moderate		
8	Maharashtra	Buldhana	Onion	Rabi	2017	Moderate		
9	Maharashtra	Chandrapur	Wheat	Rabi	2017	Good		
10	Maharashtra	Dhule	Cotton	Kharif	2017	Good		
11	Maharashtra	Gadchiroli	Tur	Kharif	2017	Poor		
12	Maharashtra	Gondia	Jawar	Rabi	2017	Good		
13	Maharashtra	Hingoli	Jawar	Rabi	2017	Moderate		
14	Maharashtra	Jalgaon	Cotton	Kharif	2017	Good		
15	Maharashtra	Kolhapur	Sugarcane	Zaid	2017	Good		
16	Maharashtra	Latur	Moong	Kharif	2017	Good		
17	Maharashtra	Mumbai city	Rice	Kharif	2017	Poor		
18	Maharashtra	Mumbai Subur	Cocoanuts	Kharif	2017	Poor		
19	Maharashtra	Nagpur	Sunflower	Rabi	2017	Good		
20	Maharashtra	Nanded	Bajra	Kharif	2017	Moderate		
21	Maharashtra	Nasik	Wheat	Rabi	2017	Moderate		
22	Maharashtra	Palghar	Cabbages	Rabi	2017	Moderate		

The dataset it contains State name, District name, Crop type, Year, and Yield rate. So, this dataset is especially of Maharashtra state with its districts. The dataset that is used needs to be pre-processed because of the existence of redundant attributes, noisy data in it. Initially, data cleaning operation is performed where the unnecessary factors are determined and are not considered for the prediction of crops. As part of the preliminary data analysis, the categorical factors are split and are assigned values as 0 and 1 based on whether the factor is present or not. These assigned values assist in further based on that particular factor. This dataset is shows which crop is growth in which district. And also shows which crop is growth in which season. So, for example in Dhule mostly cotton crop is growth. So, as per the above dataset in 2017, cotton crop's yield rate in Dhule was good.

ML Algorithm we are going to use:

Data Mining:

Data Mining is a process used to turn raw data into functional information. Data mining is a subset of marketing analytics and refers to survey an existing large dataset to unearth previously unknown patterns, relationships and anomalies that are present in the data. It gives us the ability to find completely new insights that we weren't automatically looking for unknown unknowns, if you like. Machine learning may use some data mining techniques to construct models and find patterns, so that it can make better predictions. And data mining can sometimes use machine learning techniques to produce more correct analysis [4]. So, here first we are going to do Data Mining. Data mining is the process of analyzing a large batch of data to discern trends and patterns. Data mining can be used by corporations for everything from studying about what customers are interested in or want to buy to fraud detection and spam filtering. Data mining programs break down patterns and relation in data based on what information users request or provide.

Naive Bayes Algorithm:

Naïve Bayes algorithm is a supervised learning algorithm, which is manufacture on Bayes theorem and used for solving classification problems. ... Naïve Bayes Classifier is one of the easy and most functional Classification algorithms which helps in building the fast machine learning models that can make quick predictions. Naive Bayes uses a near method to see the probability of different class based on various attributes. This algorithm is mostly used in text classification and with issue having multiple classes. Naive Bayes is a machine learning type that is used for big volumes of data, even if you are working with data that has millions of data records the recommended approach is Naive Bayes. It gives very good outcome when it comes to NLP tasks such as sentimental analysis [7]. Abstractly, Naive Bayes is a depending on probability model: given a problem example to be classified, represented by a vector representing some n features, it assigns to this instance probabilities. For each of K possible result or classes. Here we are going to use Naive bayes Machine Learning Algorithm. Because Naive bayes algorithm is accurate algorithm in ML. It is easy to

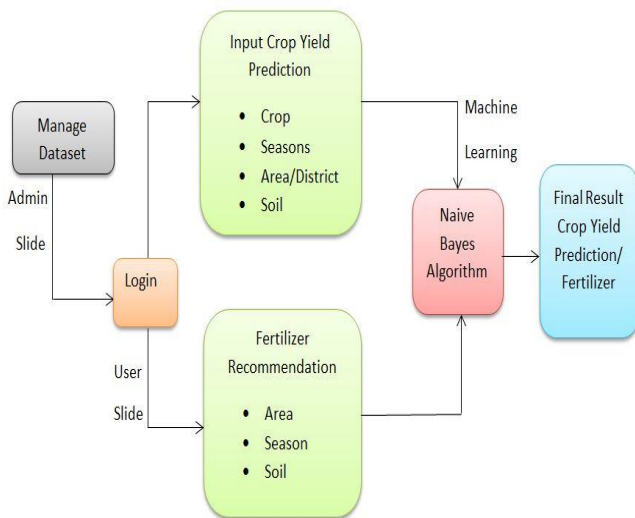
implement and it doesn't required to train the dataset again and again. Naive Bayes classifiers are a collection of grouping algorithms based on Bayes' Theorem. It is not a single algorithm but a family of algorithms where all of them is a part familiar principle.

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Probability of B occurring given evidence A has already occurred (points to P(B|A))
 Probability of A occurring (points to P(A))
 Probability of A occurring given evidence B has already occurred (points to P(A|B))
 Probability of B occurring (points to P(B))

P (A|B) is Posterior probability: Probability of hypothesis A on the observed event B. P (B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true P (A) is Prior Probability: Probability of hypothesis before observing the evidence. P (B) is Marginal Probability: Probability of verification.

Architecture:



So, in this project we will take datasets from government websites such as data.gov.in and imd.gov.in. This dataset is managed at admin side. And in the user side system, first user have to login into the system. Then, user has to give the input to a system like crop name, area or district name and season for crop yield prediction and same for fertilizer prediction. After taking input system is going to perform pre-processing on the input data. Data pre-processing is a process of preparing the raw data and making it acceptable for a machine learning model. A real-world data generally

contains sound, lost values, and maybe in an unusable format which cannot be directly used for machine learning models. Data pre-processing is required tasks for wash the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model. And after data pre-processing we get the trained dataset. After this system is going to apply the naive bayes ML algorithm to the dataset. The algorithm is contains four different modules or parts to getting the output. These four modules or parts are as follows:

1. Initial Probability:

An initial probability distribution, defined on S, define the starting state. Usually this is done by define a particular state as the starting state. Means it is the starting stage in algorithm which contains the input dataset variables.

2. Individual Probability:

Next is individual stage here some operation is going to perform on input variables. It is going to perform as: Divide the number of events by the number of possible outcomes. This will give us the probability of a single event happen. For example, In the case of rolling a 3 on a die, the number of events is 1 (there's only a single 3 on each die), and the number of outcomes is 6.

3. Conditional Probability:

Conditional probability is the probability of one event happen with some relationship to one or more other events. A conditional probability would look at these two events in connection with one another, such as the probability that it is both raining and you will need to go outside.

4. Final Probability:

In its most general case, probability can be defined numerically as the digit of desired outcomes divided by the total number of outcomes.

After all these four operation system will be produced the predicted output of it like yielding rate in Poor, moderate, good. Also we get the fertilizer prediction for the crop in rate of Poor, moderate and good. So, expected outcome of this Proposed system is its recommends the best suitable crop for particular land by considering parameter as weather conditions like humidity, rainfall, temperature & cloud. The system will use machine learning algorithm techniques to predict the yielding rate of the crop and fertilizer prediction for the crops. The proposed system will use naive bayes classifiers for crop yields prediction an as well as for fertilizer prediction. The output of the classifier will be as yielding rate in Poor, moderate, good. Also we get the fertilizer prediction for the crop in rate of Poor, moderate and good.

4. CONCLUSION

After various surveys and through analysis we can conclude that the use of machine learning algorithms will not only help farmers to get better results but also increase their revenue which for many is a matter of life and death. Currently farmers make rough estimations based on their previous experiences and plan accordingly, using ML instead

will definitely decrease the margin of error and provide them with better outputs. The proposed system will work to provide suggestions, so the final call is still left for the end user and will be very inexpensive and after some successful validations of data will provide near perfect results.

REFERENCES

- [1] Mayank Champaneri, Chaitanya Chandvidkar , Darpan Chachpara, Mansing Rathod, "Crop yield prediction using machine learning" International Journal of Science and Research ,April 2020.
- [2] Pavan Patil, Virendra Panpatil, Prof. Shrikant Kokate, "Crop Prediction System using Machine Learning Algorithms", International Research Journal of Engineering and Technology, Feb 2020.
- [3] Ramesh Medar , Shweta, Vijay S. Rajpurohit, "Crop Yield Prediction using Machine Learning Techniques", 5th International Conference for Convergence in Technology, 2019.
- [4] Trupti Bhangе, Swati Shekapure, Komal Pawar, Harshada Choudhari, "Survey Paper on Prediction of Crop yield and Suitable Crop", International Journal of Innovative Research in Science, Engineering and Technology, May 2019.
- [5] E. Manjula, S. Djodiltachoumy, "A Modal for Prediction of Crop Yield "International Journal of Computational Intelligence and Informatics, March 2017.
- [6] Nishit Jain, Amit Kumar, Sahil Garud, Vishal Pradhan, Prajakta Kulkarni, "Crop Selection Method Based on Various Environmental Factors Using Machine Learning", International Research Journal of Engineering and Technology (IRJET), Feb 2017.
- [7] Rakesh Kumar, M.P. Singh, Prabhat Kumar, J.P. Singh, "Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique", 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM),Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, T.N., India., May 2015