

The System for Maximize the Yielding Rate of Crops using Machine Learning Algorithm

Yogita Masare¹, Sneha Mahale², Manjusha Kele³, Ashvin Upadhyay⁴, Bhushan R. Nanwalkar⁵

^{1, 2, 3, 4} SVKM's Institute of Technology Dhule, India,

⁵ Department of Computer Engineering, SVKM's Institute of Technology Dhule, India,

Abstract:- As we are aware of the fact that, most of Indians have agriculture as their occupation. Farmers usually have the mind-set of planting the same crop, using more fertilizers and following the public choice. By looking at the past few years, there have been significant developments in how machine learning can be used in various industries and research. So, we have planned to create a system where machine learning can be used in agriculture for the betterment of farmers.

India is an Agricultural Country and its economy largely based upon crop productivity. So we can say that agriculture can be pillar of all business in our country. Selecting of always crop is very important in the agriculture planning. Many researchers studied guess of yield rate of crop, guess of weather, soil categorizing and crop classification for agriculture planning using machine learning techniques. Many changes are required in the agriculture department to improve changes in our Indian economy. We can improve agriculture by using machine learning system which are applied simply on farming sector. Along with all advances in the machines and technologies used in farming, functional 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 problems. This enhances our Indian wealth by maximizing the yield rate of crop production.

Key words: Agriculture, Crop Selection Method, Economy.

1. INTRODUCTION

Agriculture plays a very important role where economic growth of a country like India is considered. The main aim of agricultural planning is to achieve maximum yield rate of crops by using little number of land resources. Many machine learning algorithms can help in raising the producing of crop yield rate. Whenever there is loss in critical situation, we can apply crop selecting method and reduce the Losses. And it can be used to acquire crop yield rate in favourable conditions. This Maximizing of yield rate helps in upgrade countries economy. We have some other factors that impact the crop yield rate. They are fertilizer quality and crop Selection. Selection of crops depends upon two task that are favourable and critical conditions. Many researches are carried out to improve agricultural setting. The aim is to get the biggest yield of crops. Many categorizing systems are also applied to get maximum yield of crops. Machine learning can be used to better the yield rate of crops.

The method of crop Selection is bother to improve crop production. The construction 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, salt or peaty. Soil composition can be copper, potassium, sulphate, nitrogen, manganese, iron, calcium, ph. value or carbon and different methods of harvesting. Many parameters are used for different Crops to do separate predictions. These prediction models can be studied by using analyzer. These predictions are classified as two types. One is traditional statistic method and other is machine teaching techniques. Traditional method helps in predicting single sample spaces. And machine learning methods helps in predicting different predictions. We need not to consider the design of data models in traditional method where as we need to consider the structure of data models in machine learning methods.

2. LITERATURE REVIEW

A. Crop Prediction System Using Machine Learning Algorithms^[1]

In this paper, the author 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.

B. Crop yield Prediction using Machine Learning^[2]

This paper builds 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.

C. Crop yield Prediction Using Machine Learning Techniques^[3]

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.

D. Crop Selection Method Based on Various Environmental Factors Using Machine Learning^[4]

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 crops which can be grown throughout the season.

E. A Modal for Prediction of Crop Yield^[5]

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 use 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.

F. Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique^[6]

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. PROPOSED SYSTEM

All we know that farming is the backbone of the Indian economy. But now a days the modern people don't have awareness about the growing of crops at the right time and at the right place. because of these cultivating techniques the seasonal climatic condition is also being changed against the basic's 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 assist of those algorithms we can predict yield also.

We focus on main four objectives:

- 1) To use Machine Learning Algorithm to predict crop yield
- 2) To provide easy to use interface
- 3) To increase the validity of crop yield prediction
- 4) To analyses different climatic framework 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.

3.1. Crop yield prediction module:

Crop Yield Prediction is the methodology to predict the yield of the crop using different parameters like temperature, rainfall, pH values, EC, CEC, OC, SAR, pesticides, etc. Crop yield is a measurement of the number of agricultural production crop. There are four main elements 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 investing 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 total invested.

There are a number of factors which are likely to have provide to sustained yield gains: fertilizer request, 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 name and season for crop yield prediction. After this system is going to apply Machine learning algorithm to the input data and it will be produced the output of it like yielding rate in Poor, moderate, good for the crop.

3.2. Fertilizer prediction:

Plants need to be fertilized because most soil does not manufacture the required nutrients required for optimum growth. It looks nutrients from the soil to build that plant material. By fertilizing your park, you reload lost supplements and make sure that this year’s plants have the food they need to flourish. Excess fertilizer alters the soil by generate too high of a salt absorption, and this a hurt beneficial soil microorganism. Over fertilization can lead to unexpected 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 insert to a system like crop name, area or district name and season for fertilizer prediction. After this system is going to apply Machine Learning 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 be good or bad god the crop.

Dataset:

SrNo	Crop	Season	Soil	Temperature	Rain	PH	EC	CEC	OC	SAR	Output
1	Wheat	Winter	SANDY	Low	Low	Medium	Low	High	Low	Medium	Low
2	Wheat	Summer	SANDY	Low	Low	Low	Low	Low	Low	Low	Low
3	Wheat	Rainy	SANDY	Low	Low	Low	Low	Low	Low	Low	Low
4	Wheat	Winter	SANDY	Low	Low	Low	Low	Low	Low	Medium	Low
5	Wheat	Winter	SANDY	Low	Low	Low	Low	Low	High	Medium	Low
6	Wheat	Winter	SANDY	Low	Low	Low	Low	High	High	Medium	Low
7	Wheat	Winter	SANDY	Low	Low	Low	Medium	High	High	Medium	Low
8	Wheat	Winter	SANDY	Low	Low	High	Medium	High	High	Medium	Low

Naive Bayes Algorithm:

Input:

Training Dataset D
 P= (p1, p2, p3... pn) Parameters in dataset
 T= (t1, t2, t3 ...tn) Testing tuple

Output:

O= (o1, o2, o3) output variables in dataset

Steps:

- 1) Read the dataset Training
- 2) Calculate frequency count for output variables

- 3) Calculate individual probability for output variables
- 4) Calculate conditional probability for output against each input variables in test tuple
- 5) Calculate final probability for all output (Step 3 probability * step 4 probability)
- 6) Find the greatest probability for against output variables
- 7) Return output variable with maximum probability

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

Fig 1. Naive Bayes Formula

Data Mining:

Data Mining is a process used to turn sensitive data into functional information. Data mining is a subset of marketing analytics and refers to survey an existing large dataset to unearth previously unspecified 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 unknowns, if you like. Machine learning may use some data mining techniques to build models and find design, so that it can make better predictions. And data mining can sometimes use machine learning techniques to generate more correct analysis. 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 sample. Data mining can be used by firm for everything from studying about what customers are interested in or want to buy to fraud detection and spam filtering. Data mining programs crack 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 grouping problems. Naïve Bayes Classifier is one of the uncomplicated and most functional Classification algorithms which help 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 many attributes. This algorithm is mostly used in text classification and with issue having multiple categories. Naive Bayes is a machine learning type that is used for big capacity 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 investigation. Abstractly, Naive Bayes is a depending on probability model: given a problem

case to be classified, represented by a vector representing some n features, it assigns to this instance probabilities. For each of K possible result or category. Here we are going to use Naive bayes Machine Learning Algorithm. Because Naive bayes algorithm is accurate algorithm in Machine Learning. It is easy to implement and it doesn't require to train the dataset again and again. Naive Bayes classifiers are a collection of grouping algorithms based on Bayes' Theorem. It is not a one algorithm but a family of algorithms where all of them is a part familiar principle.

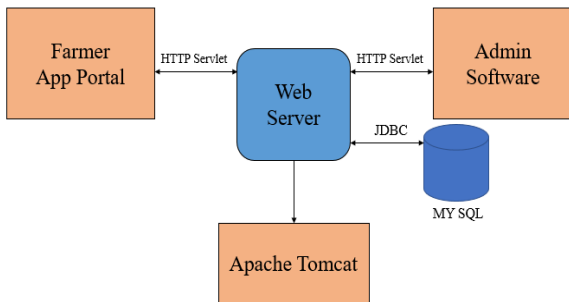


Fig 2. System Architecture

So, above is the over system architecture. It is consisting three modules and that are, Web Server, Admin Software and Farmer App/Portal. The Web Server is manages the database. So here we use the apache Tomcat for web server. Basically, Apache Tomcat is a free and open-source implementation of the Java Servlet, Java Server pages, Java Expression Language and WebSocket technologies. Tomcat provides a pure Java 'HTTP web server' environment in which Java code can run. And in the admin software, admin is going to manage the dataset. And in the user side i.e., Farmer App or Portal, users are going to provide an input to the system and system is going to perform the operation on the input parameters and then it is going to give the appropriate output to the farmer.

Admin Side:

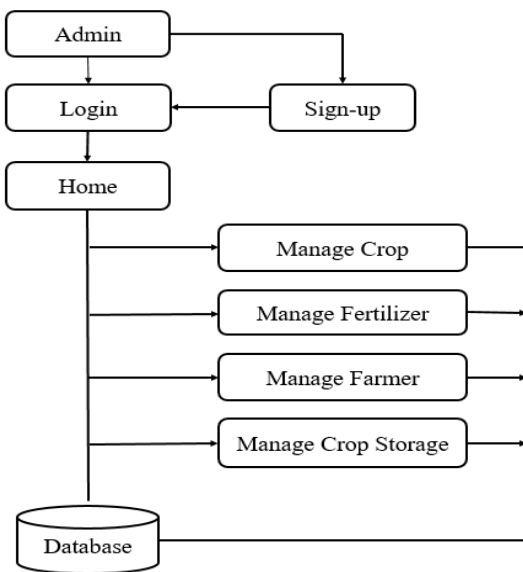


Fig 3. Admin Side Software

User Side:

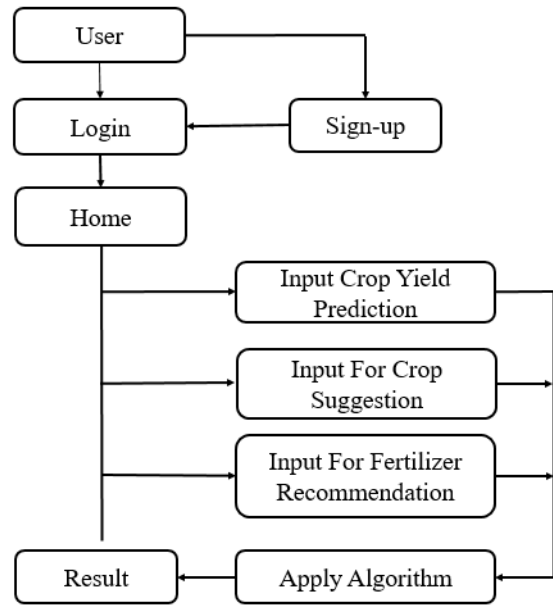


Fig 4. User Side Portal

So, in this project we will take datafile from government websites such as data.gov.in and imd.gov.in. This datafile is manages at admin side. And in the user side system, first user has to login into the structure. If he or she having not an account he or she have to sign up and make their account and then they can be login into the system. Then, user have to give the input to a system i.e., season name for crop yield prediction. Again, season name and crop name as input for fertilizer prediction. After taking input system is going to perform pre-processor the input data. Data pre-processing is a process of produce the raw data and making it acceptable for a machine learning model. A real-world data generally contains sound, old values, and maybe in an unusable format which cannot be directly used for machine learning models. Data pre-processing is required tasks for clean 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-processor we get the trained dataset. After this system is going to apply the naive bayes supervise Machine learning algorithm to the dataset. The algorithm is contains four different modules or parts to acquire the output. These four modules or parts are as follows:

A. Initial Probability:

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

B. Individual Probability:

Next is individual stage here some operation is going to execute on input variables. It is going to perform as: Divide the number of events by the figure of possible outcomes. This will give us the expectation of a single event happen.

For example, In the case of rolling a 4 on a die, the number of Events is 1 and the number of outcomes is 8.

C. Conditional Probability:

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

D. Final Probability:

In its most general case, probability can be defined numerically as the no of desired outcomes divided by the total number of outcomes.

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

4. RESULT AND DISCUSSION

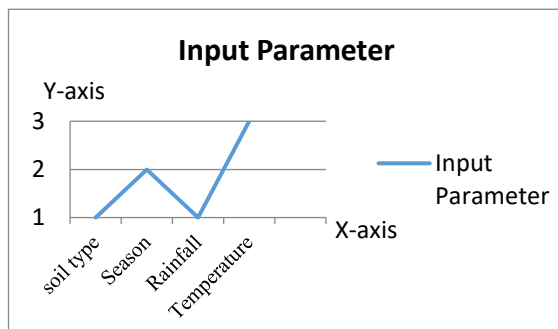


Fig: Case 1-Graph of Input parameter

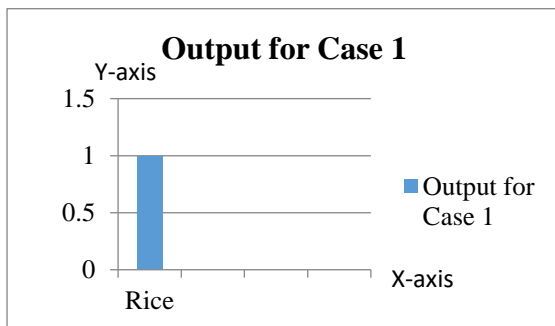


Fig: Output for Case 1

As per shown in the case1 the point on axis 1 is for low, 2 is used for medium and 3 is used for High. After entering this input parameters system is going to perform operation on it for giving the output. So, as per shown in figure of output it shows that crop rice is the output of the input of case 1. That means, if the soil type is clay (1is used for Clay), season is winter, rain fall is medium and temperature is high then we can go for Rice crop for getting better yield.

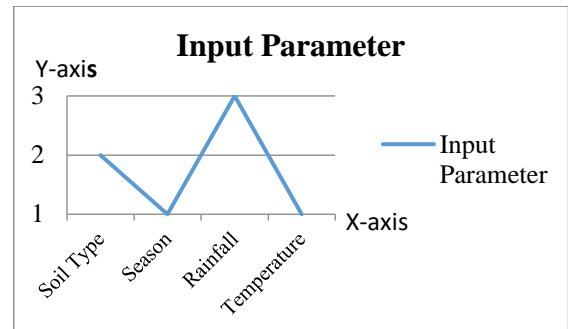


Fig: Case 2- Graph of Input Parameters

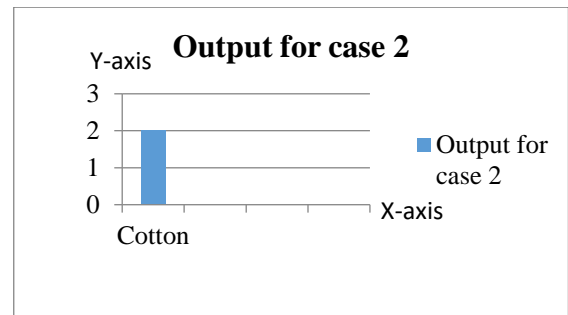


Fig: Output for Case 2

As per shown in the case 2 the point on axis 1 is for low, 2 is used for medium and 3 is used for High. After entering this input parameters system is going to perform operation on it for giving the output. So, as per shown in figure of output it shows that crop rice is the output of the input of case 2. That means, if the soil type is sandy (2 is used for Clay), season is summer, rain fall is high and temperature is low then we can go for Cotton crop for getting better yield.

CONCLUSION

After going through many surveys and through analysis we can conclude that the use of no of various 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. This proposed system will work to provide suggestions, which definitely going to helps farmers to get more yield and better crops.

REFERENCES

- [1] Mayank Champaneri, ChaitanyaChandvidkar, DarpanChachpara, MansingRathod, "Crop yield prediction using machine learning" *International Journal of Science and Research*, April 2020.
- [2] Pavan Patil, VirendraPanpatil, Prof.ShrikantKokate, "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] TruptiBhange, Swati Shekapure, Komal Pawar, HarshadaChoudhari, "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, SahilGarud, 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.*