



CROP YIELD PREDICTION USING MACHINE LEARNING AND CLOUD COMPUTING

Disha A

MCA Scholar, School of CS & IT, Dept. of MCA Jain(Deemed-to-be) University, Bangalore

ABSTRACT

Agriculture is the field which assumes a significant part in improving our nations economy. India is an agrarian country and its economy is generally founded on crop efficiency. Choosing of each yield is vital in the farming arranging. The determination of yields will rely on the various boundaries, for example, market value, creation rate and the distinctive government approaches. Numerous progressions are needed in the farming field to improve changes in our Indian economy. We can improve crop yield by utilizing AI procedures which are applied effectively on cultivating area.

Stretching out web based business to the cultivating local area is a significant advance in the computerized change of India. For Mr. Gupta, it addresses the democratization of internet business. Today, ranchers have the choice to purchase online as opposed to going through hours making a trip to and from actual retail locations. "We are digitizing agribusiness in India with AWS—engaging ranchers with internet business availability. Likewise, by exhibiting recordings on successful harvest cultivating, ranchers have had the option to build their yield yields by in excess of 25%," Mr. Gupta says.

With the site running on the AWS Cloud, improvement time is a lot more limited. Mr. Gupta remarks, "With AWS, our IT activities productivity has improved by at any rate 80% while creating cost investment funds of in excess of 50%, on the grounds that we can choose, convey, and oversee AWS assets effortlessly. There is no acquirement cycle or need to devote capital spending plans to projects. Our month to month AWS charging is not even close to the expense of power for an on-premises climate of a comparative size."

KEYWORDS: *Indian Agriculture, Machine Learning Techniques, Crop selection method.*

II. INTRODUCTION

The main goal 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 improving the production of crop yield rate.

Whenever there is loss in unfavourable 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 maximising of yield rate helps in improving countries economy.

We have some of the factors that influence the crop yield rate. They are seed quality and crop selection. We need test the quality of the seeds before sowing. As we know that good quality of seeds helps in getting more yield rate. And selection of crops depends upon two things that is favourable and unfavourable conditions. This can also be improved by using hybridisation methods. Many researches are carried out to improve agricultural planning. The goal is to get the maximum yield of crops. Many

classification methods are also applied to get maximum yield of crops.

Machine learning techniques can be used to improve the yield rate of crops. The method of crop selection is applied to improve crop production.

For the better yield we need to consider soil type and soil fertility things and also one of the major factors rainfall and groundwater availability if it is dry land it is better to go for cash crops and if is wetland it is better to go for wheat and sugarcane.

There are 15 agro-climatic regions in India these regions are divided on the bases of a type of the land. Each agro-climatic regions can grow some specific crops. Based on that we need to suggest the farmer that which crop is best among those crops which belong to those climatic regions.

Achieving the maximum crop at minimum yield is the ultimate Aim of the project. Early detection of problems and management of that problems can help the farmers for better crop yield. Crop yield prediction is the important research which helps to secure food. For the better understanding of the crop



yield, we need to study of the huge data with the help of machine learning algorithm so it will give the accurate yield for that crop and suggest the farmer for a better crop.

Improving the quantity of the crop is the key goal of precision agriculture means obtaining a better understanding of the crop using the information technology methods. The main goal of precision agriculture is profitability and sustainability

III.OBJECTIVE OF THE PROJECT

The proposed structure targets predicting or expecting the gather yield by learning the past data of the developing area. By thinking about various variables, for instance, soil conditions, precipitation, temperature, yield and various substances the system manufactures an expecting a model using AI techniques. Here we use unmistakable AI strategies such sporadic forest, Polynomial Regression, Decision Tree. Execution is surveyed subject to expected exactness.

IV. LITERATURE SURVEY

In [1] J.P. Singh, Rakesh Kumar, M.P. Singh and Prabhat Kumar, have presumed that this paper helps in improving the yield pace of harvests by applying grouping strategies and contrasting the boundaries. We can likewise do dissecting and expectation of harvests utilizing baysian calculations. The calculations utilized are Bayesian calculation, K-implies Algorithm, Clustering Algorithm, Support Vector Machine. The burden is that there could be no appropriate exactness and execution.

In [2] the creators Subhadra Mishra, Debahuti Mishra and Gour Hari Santra, have inferred that this is a high level explored field and is relied upon to fill later on. The joining of software engineering with horticulture helps in determining rural yields. This technique additionally helps in giving data of harvests and how to expand yield rate. The calculations utilized are Artificial neural organizations, Decision Tree Algorithms, Regression investigation. The determination is clear approach isn't determined.

Monali Paul, Santosh K. Vishwakarma, Ashok Verma[3] In request to foresee the yielding of the harvests, the harvests are broke down and dependent on investigation they are arranged. This categorisation is done dependent on information mining calculations. This paper gives knowledge into different arrangement rules like Naive Bayes, K-Nearest Neighbor. Utilizing this paper, we dissected the grouping rules and recognized which will be suitable for informational index which we will use in our task.

V. PROBLEM STATEMENT

Early forecast of harvest yield is significant for arranging and taking different strategy choices. Numerous nations utilize the customary procedure of information assortment for crop checking and yield forecast dependent on ground based visits and reports. These techniques are abstract, exorbitant and tedious. The normal issue in existing harvest yield forecast techniques are given underneath,

The main issue of existing harvest yield expectation strategy is exactness and tedious issue. In existing time arrangement crop yield forecast strategy doesn't respond to varieties that happen for cycles and occasional impacts. Needs broad data to create and test the model and furthermore accessible data in farming is inadequate and fragmented in existing reenactment model. Restricted investigations have been made in crop yield forecast utilizing existing choice tree strategy. Expectation mistake esteem likewise significant issue in crop yield forecast or assessment strategies. These are the primary disadvantages of different existing works, which inspire us to do this examination on crop yield expectation.

VI. EXISTING SYSTEM

An agro-put together nation depends with respect to farming for its financial development. At the point when a populace of the nation builds reliance on horticulture additionally increments and ensuing monetary development of the nation is influenced.

In the present circumstance, the harvest yield rate assumes a huge part in the financial development of the country. Along these lines, there is a need to expand crop yield rate. Some organic methodologies (for example seed nature of the yield, crop hybridization, solid pesticides) and some substance draws near (for example utilization of manure, urea, potash) are done to tackle this issue.

Notwithstanding these methodologies, a harvest sequencing procedure is needed to improve the net yield pace of the harvest over the season. One of existing framework we distinguished is Crop Selection Method (CSM) to accomplish a net yield pace of harvests over the season. We have taken illustration of CSM to show how it helps ranchers in accomplishing more yield Crop can be named:

a) Seasonal yields—harvests can be planted during a season. for example wheat, cotton.

b) Whole year crops—harvests can be planted during the whole year. for example vegetable, paddy, Toor.

c) Short time manor crops—crops that set aside a short effort for developing. for example potato, vegetables, proportion. d) Long-time manor crops—These yields set aside a long effort for developing. for example sugarcane, Onion. A mix of these harvests can be chosen in a grouping depend-



ent on yield rate each day. Represents successions of harvests with aggregate yield rate over the season. CSM strategy, appeared in may improve the net yield pace of harvests utilizing the restricted land asset and furthermore expands re-use of use of the land.

Essentially, in crop choice strategy utilizes method where it suggests diverse arrangement of yields for same territory throughout the long term. There are different alternatives are accessible to choose for ranchers.

They can pick one of the alternatives and notice the outcomes. The mix which will give high return for same territory is produced as yield for that space. In this manner CSM strategy attempts to anticipate the reasonable yields for given region. Cultivating Systems in India are deliberately used, as indicated by the areas where they are generally appropriate.

VII. PROPOSED SYSTEM

A. Acquisition of training dataset:

The exactness of an AI calculation may rely upon the quantity of boundaries utilized and to the degree of rightness of the dataset. Our dataset contains the N, P, K, and pH upsides of various types of soils as properties and it likewise contains the comparing crops that can be filled in that dirt as mark. Accordingly, by utilizing a fitting AI calculation we can prepare the dataset to anticipate the most reasonable yield that can be developed under the given information boundaries.

B. Data preprocessing:

Information preprocessing is the subsequent advance and it contains two stages. Unique dataset can contain heaps of missing qualities so at first all these ought to be taken out. Missing qualities are signified by a dab in the dataset and their essence can decay the worth of whole information and it can lessen the exhibition.

Thus, to tackle this difficult we supplant these qualities with enormous negative qualities which will be treated as anomalies by the model. Producing the class marks is the subsequent advance. Since we are utilizing a regulated learning strategy, for every section in the dataset there ought to be a class name which is made during the preprocessing step.

VIII. METHODOLOGIES

[1] Regression analysis

Regression analysis a measurable method that is utilized to foresee the qualities from the normal yield amount when the yield amount is consistent.

[2] Linear Regression

Linear regression is a linear method to build a relationship between two variables which are continuous – one variable which is independent and denoted by X and another variable which is dependent and denoted by Y. This is used for prediction which is based on statistical methods. The graph of linear regression classification is a straight line. The formula for calculating the linear regression is-

$$y' = b + ax$$

calculate the slope by calculating b.

[3] Non-Linear Regression

Nonlinear regression is a regression technique which does not depend on one variable but depends on multiple variables or predictor variables. The graph of this nonlinear regression is a curve where all the variables are dependent on each other. This model is used to calculate the sum of squares which is the dispersion of data points. This method is based on trial and error technique and needs many assumptions. The trial and error method is used to minimise the value of sum of squares to make the data point fit in the data set.

[4] Dataset Description

Generally researchers used .csv files of agriculture dataset for crop yield prediction. The dataset is supervised learning. It consists of different attributes like County Name, State, humidity, temperature, NDVI, wind Speed, yield

Crop yield is the measure of crop produced per area of land. It's an important metric to understand because it helps us understand food security and also explains why your tomatoes can cost more one year and then less the following year.

IX. CONCLUSION

This project is undertaken using machine learning and evaluates the performance by using Random forest, Polynomial Regression and Decision Tree algorithms. In our proposed model among all the three algorithm Random forest gives the better yield prediction as compared to other algorithms. Along with random forest, Polynomial Regression, Decision Tree model classify the output that shows improvements in dataset. So we analysed that proposed model has got more efficiency than the existing model for finding crop yield.

REFERENCES

1. J.iajun Zong, Quanyin Zhu, "Apply Grey Prediction in the Agriculture Production Price", Fourth International Conference on Multimedia Information Networking and Security, 2012.



2. Michael gurstein, "a decision support system to assist the rural poor in Bangladesh", *IEEE TECHNOLOGY AND SOCIETY MAGAZINES*, September 2013.
3. Monali Paul, Santhosh K. Vishwakarma, Ashok Verma, "Prediction of Crop Yield using Data Mining Approach" *Computational Intelligence and Communication Networks(CICN)*, International Conference 12-14 Dec. 2015.
4. Tng Zhang, "Solving large scale linear prediction algorithm", *proceedings of the twenty-first international conference on Machine Learning* Aakunuri Manjula, G. Narsimha, "XCYPF: A Flexible and Extensible Framework for Agricultural Crop Yield Prediction", *IEEE Sponsored 9th ISCO, 2015*.
5. Aakunuri Manjula, G. Narsimha, "XCYPF: A Flexible and Extensible Framework for Agricultural Crop Yield Prediction", *IEEE Sponsored 9th ISCO, 2015*.
6. D. Ramesh, B. Vishnu Vardhan, "Analysis Of Crop Yield Prediction Using Data Mining Techniques", 2015
7. A.A. Raorane, R. V. Kulkarni, "Data Mining: An effective tool for yield estimation in the agricultural sector"< *IJETTCS*, vol. 1, no. 2, pp. 75-79, 2012.