4th National Conference on Green Technology and Science for Sustainable Development
© 2023 IJSRST | Volume 10 | Issue 10 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
International Journal of Scientific Research in Science and Technology
doi : https://doi.org/10.32628/IJSRST231010147

Forecasting of Optimal Crop Yield through Data Mining Algorithm

Natvar S Patel, Dr. Bhavesh M Patel
Research Scholar, Department of Computer Science, HNGU University, Patan, Gujarat, India
I/C Head of Department, Department of Computer Science, HNGU University, Patan, Gujarat, India

ABSTRACT

The initial stage of analytics is descriptive analytics. It is a method via which we can learn about the past. We are aware that the past serves as the finest indicator of the upcoming. That is study, Descriptive analytics are used by us to estimate crop yields effectively for the production of sugarcane in agriculture. Three datasets—the soil, rainfall, and yield datasets—are used in this paper. In order to determine the real estimated cost and the accuracy of various strategies, we create a composite dataset and then apply a number of supervised techniques to it. Three supervised techniques—crop prediction, provision direction, and LSS method—are employed in this research. It is a comparative study that reveals the mistake rate and accuracy of training the suggested model. The training model should be more accurate and have a lower error rate than possible. Additionally, the suggested model can provide the real cost of the projected crop production and can be classified as LOW, MID, or HIGH.

Keywords: Crop forecasting, supply planning, and LSS methodology, Data Analytics, Agriculture analytics.

I. INTRODUCTION

A large portion of India's economy depends on agriculture. One of the country's key industrial sectors and is essential for rural sustainability. India's agricultural output is falling as a result of various issues, including changing weather patterns, declining water levels, overuse of pesticides, etc. Regarding agriculture data, we used descriptive analytics to determine the level of production. The major goal of this research project is to offer a methodology that can be used to effectively perform descriptive analytics on agricultural yield output. A few studies have looked at crop prediction using historical meteorological and production data, despite the fact that several studies have disclosed statistical information on Indian agriculture.

We discovered that numerous models, including Partial least squares, principal component regression, adaptive forecasting, etc., are in use. However, these models share the fact that they are either based on categorization as well as regression. We are currently creating a programme that is monitored dependent. Additionally, this will function as a mixed method, which includes both regression and classification techniques.

In our project, crop yields will be classified according to yield productivity, with the classes being labeled as low, mid, and high. To determine the real crop yield estimated cost, a productivity range will be specified and regression analysis will be done. This is why this system was created. Crop yield forecast models are created for

projecting yield well in advance of the actual harvest of the crops inspired by research of crop conditions. Plants yields are actually operationally forecasted in the nation using empirical statistical models that employ correlation and regression techniques.

The models take into account both technological trends and meteorological elements at various stages of crop growth. Additionally, if we ever create a comprehensive farmer recommender system, the results of this research will be useful. Since descriptive analytics, the cornerstone of all recommender systems, is being performed here.

II. The Suggested System

(1) The following list of factors includes a number that significantly influence agricultural production.
(2) Plant Production Variation with Rainfall
(3) A key element in the production of agriculture is rainfall. Rainfall variations may have an impact on crop productivity. A high rainfall rate may contribute to low productivity. There shouldn't be too much rain. The optimal rainfall range is between 300 and 600 millimetres, which may result in exceptionally high or ordinary output.
(4) Changes in the Humidity Factor
(5) The output of agricultural goods is significantly influenced by the metrological characteristics. One of the most crucial metrological parameters is the humidity. Any crop can grow better when the conditions are humid.
(6) Climate Change's Effect on Agriculture

A crucial component of agriculture is the climate. The agricultural system is significantly impacted by climatic changes. The amount of precipitation is crucial to agriculture. The patterns of rainfall have changed as a result of climate change. This leads the majority of crops to require water. These crops are costing more than they did in the beginning. Because of increased tube well and other resource utilisation due to a lack of rainfall, the irrigation procedure has grown more expensive.

• Effects of Climate Change Globally

Viewpoints that are explicit and revocable can be used to arrange the effects of elevation change. Because of the global temperature rise, some overall impacts can be encouraged, such as less severe winters and more vegetation in high altitude regions. However, compared to the superior effects, the antagonistic (negative) effects clearly come out on top. Some of the negative consequences brought on by elevation change, particularly when combined with sanctifying through water and an Asian aesthetic, can be structured as follows. And the following are some of the causes:

Global warming has an impact on the environment.

• Pollution and excessive use of electronic devices like refrigerators and air conditioners.
• The majority of scientific experiments include nuclear power.
• Pollution in the lakes, which is a major contributor to sea pollution.
• The indiscriminate use of fuel on a global scale.
• There is a rapid melting of the glaciers. And this is the primary reason why the water level rose.

Crop Prediction Model

Crop prediction model , In essence, it combines moving average with autoregressive models. A model that is auto-regressive and in which the response variable is determined by its prior values. The error rate is included in the moving average's numbers, though. The Crop Prediction Model works best when there is non-stationary, which indicates that there are trends in the series but they are not all equal, which causes variance to be different. Differentiating the equation of interval values, we attempt get Higher and lower bound to solve difficulties.

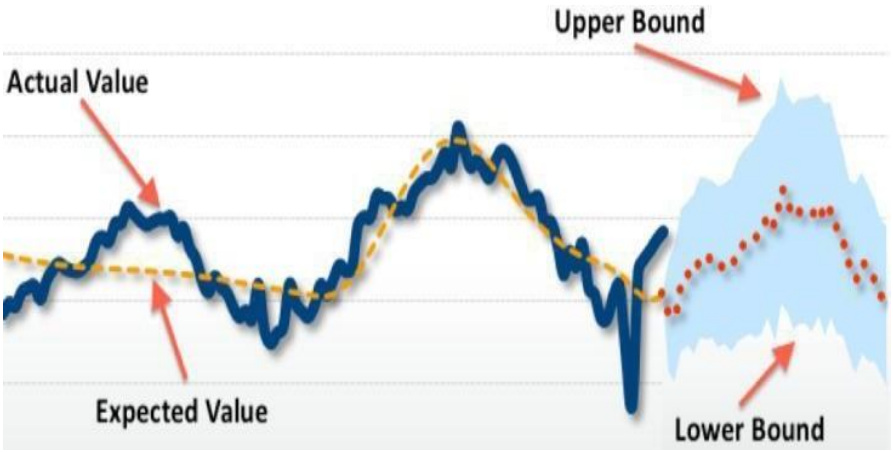

Fig 1: A model for predicting crop failure

Problem Formulation

Several climatic elements have an impact on agricultural production. Examples include Temperature, humidity, wind speed, and moisture are examples of metrological characteristics. Rainfall parameters include precipitation specific to the area rainfall, water supply, etc., as well as factors related to soil including acidity, carbon in the soil, calcium, and cellulose. And because of the continued effects of climate change, everything is in disarray.

Indian farmers continue to employ the traditional techniques they inherited from their ancestors. The problem is that the environment was fairly wholesome when everything proceeded according to plan. But most things have changed recently as a result of climate change and countless additional elements. The absence of consistent rainfall is India's main issue with agriculture. Although excessive humidity has become a problem, it is still required for crops. Rabi crops have been severely impacted by the winter season. Since a few years, wintertime precipitation has been higher than normal.

To solve the aforementioned problems, a system that can uncover hidden outcomes, patterns, and insights must be created. The farmer can foresee which crop should be sown in order to maximise profit. In the proposed system, we are using data analytics methods to analyse datasets based on agricultural productivity and uncover insights that can aid farmers in making decisions.

In this research, a system built on descriptive analytics is suggested. how farmers can be informed about

current events and future developments. As a result, we gather various data from agricultural productivity, Here, you can construct the relevant datasets for soil, rainfall, and other data.

In this strategy, the crop yield production cost and the appropriate production class are provided by the model once it has been trained using supervised learning.

Dataset gathering

We gather data from numerous sources and create datasets at this step. Additionally, these datasets are employed in both illustrative and analytical analytics. Abstracts can be found online in a variety of places, including Data.gov.in and indiastat.org. We'll employ yearly summaries of agricultural data for at least ten years. These datasets often permit time series with chaotic behaviour. integrated the primary and necessary abstracts for sugarcane (data were gathered for at least 10 years). agriculture information.

Elements	Very low	Low	Medium	High	Very high
pH	<5.0	5.1 -6.5	6.6 -7.5	7.6 -8.0	>8.0
Organic carbon(OC) in %	<0.25	0.50- 0	0.51- 0.75	0.76 - 1.00	>1.00
Nitrogen (N) in kg/ha	<150	151 - 250	251 -400	401 -600	>600
Phosphorus (P) in kg/ha	<5	06-10	11-20	21 -40	>40
Potassium (K) in kg/ha	<200	201 -250	251 -400	401 -600	>600
Zinc (Zn) in mg/kg	<0.30	0.31 - 0.60	0.61- 1.20	>1.20	Not Defined
Iron (Fe) in mg/kg	Not Define	<4.5 0	4.51 -9.0	>9. 0	Not Defined
Copper (Cu) in mg/kg	Not Define d	<0.2 0	0.21- 0.40	>0. 40	Not Defined
Manganese (Mn) in	<1.0	1.0 - 2.0	02-Apr	>4. 0	Not Defined
Sulphur (S) in kg/ha	<10	11-20	21 -30	31 -40	>40

Table 1: Land Factors for Agriculture Crop Estimation

Crop Prediction Method: The creation of a crop anticipation model will be the main emphasis of this step; the model will create a deterioration archetype to predict a harvest of the yield in the near future due to many characteristics. The following multi-stage process begins with the addition of modified statistical features. The advanced range of modified measurements that aid in recognising the arrangement's motion by anticipating fresh information based on the unbiased values will make up these altered appearances, which won't be specifically related to the crop anticipated acreage. The goal at this moment is to assemble as many affected appearances as you can in order to supply enough data for the method's later steps.

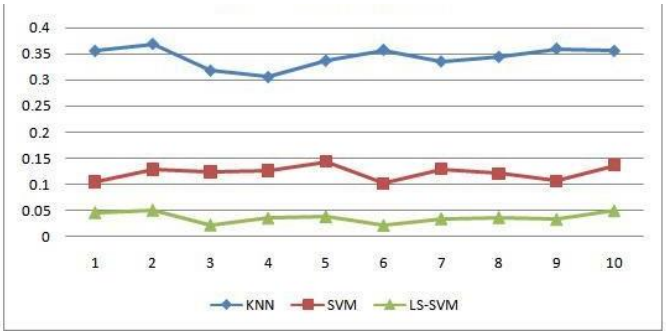

Fig 3 : LSS All classifiers for estimating crop yield have errors

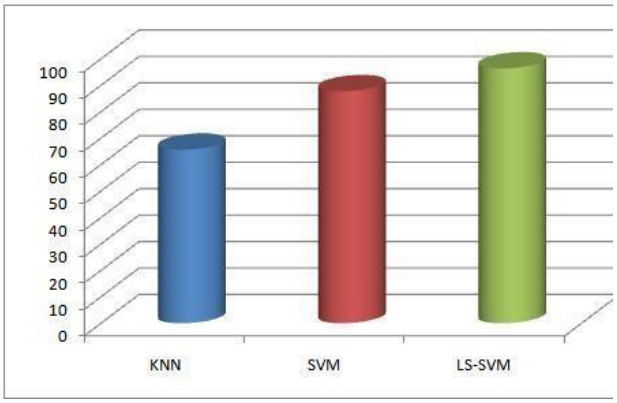

Fig 4: Agriculture Yield Prediction Classifiers' Overall Efficiency

Summary and Future Directions

The report proposes the use agricultural industry's use of descriptive analytics. That was done effort contains information for the potential use of sugarcane crop datasets for data analytics. The data on groundwater, the data on rainfall, and the data on productivity are the three available datasets. These datasets, which were collected from the agriculture domain, contain a variety of factors that may be used to classify the data into different groups and determine the health of the plants.

This system is capable of doing both regression and classification. The classification stage divides the data The regression stage calculates the real cost of yield generation and divides it into three categories (low, mid, and high). In this job essentially offers a comparative analysis of various algorithms when we use them to train datasets. It also displays the The average square error for the sample data's cross-validation phase and the accuracy of each algorithm when used to train datasets. This effort transcends domain boundaries. It implies that we can develop systems for other industries, such as medicine, retail, and product comparison. The datasets just need to be passed through this system, but they must be in a uniform format.

This scientific project has potential for advancement. We can create a system of recommendations for distributors and producers of agricultural goods for farmers. the best crop to grow at the best moment to maximise your output. Structured dataset management is the purpose of the system. Future plans call for the implementation of data independent systems. It implies that regardless of the form of data, our system ought to work effectively.

REFERENCES

(1) "CropSyst, a cropping systems simulation model: water/nitrogen budgets and crop yield," Agricultural Systems, 1994, vol. 46(3), pp. 335-359. C. O. Stockle, S. A. Martin, and G. S. Campbell.

(2) "Big Data in Precision Agriculture: Weather Forecasting for Future Farming," in First International Conference on Next Generation Computing Technologies, 2015, pp. 744–750. Bendre, M. R., Thool, R. C., and Thool, V. R.

(3) Learning dynamics of pesticide abuse by data mining: Abdullah, A., Brobst, S., Pervaiz, I., Umer, and A. Nisar, 2004. Australian Workshop on Data Mining and Web Intelligence Proceedings, January, New Zealand.

(4) C. O. Stockle., S. A. Martin and G. S. Campbell, "CropSyst, a cropping systems simulation model: water/nitrogen budgets and crop yield," Agricultural Systems, 1994, vol. 46(3), pp. 335-359.

(5) "Big Data in Precision Agriculture: Weather Forecasting for Future Farming," in First International Conference on Next Generation Computing Technologies, 2015, pp. 744–750. Bendre, M. R., Thool, R. C., and Thool, V. R.

(6) A.Nisar, A. Abdullah, S. Brobst, I. Pervaiz, and M. Umer. "Learning dynamics of pesticide abuse through data mining" (2004). Australian Workshop on Data Mining and Web Intelligence Proceedings, January, New Zealand.

Cite this Article

Natvar S Patel, Dr. Bhavesh M Patel, "Forecasting of Optimal Crop Yield through Data Mining Algorithm", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 10 Issue 10, pp. 352-357, May-June 2023. Available at doi : https://doi.org/10.32628/IJSRST231010147
Journal URL : https://ijsrst.com/IJSRST231010147