Volume 9, Issue 5, May – 2024 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY215

Method and System for Predicting Crop Yields and Recommending Fertilizers using Machine Learning Algorithms

[1]Spoorthi A. Hunshal; [2]Sanjana R.; [3]Himani Pitta; [4]Shiva Kumar R. Naik

Abstract:- In the recent years ensuring food security plays a major role in the agricultural sector and contributing towards the nations growth. This paper presents a communal system that makes use of elegant machine learning techniques and models to forecast accurate yield of the selected crop and advocates the felicitous fertiliser. By exploiting the agricultural datasets, our system employs the Extra Trees Regressor trees which helps for the prediction of the yield and to analyse the most recommendable fertilizer it makes use of the Gaussian Naïve Bayes algorithm. This dormant system provides the us with powerful insights. Our sight is to reshape the conventional agricultural practices with the help of these powerful insights to redefine the farming practices and to increase the productivity, therefore ensuring the legitimate agricultural practices.

Keywords:- *Crop Yield Prediction, Fertilizer Recommendation, Machine Learning Algorithms, Extra Trees Regressor (ETR), Gaussian Naïve Bayes (GNB).*

I. INTRODUCTION

Agriculture plays an indispensable role in the economies of developing nations like India, serving as the main support for comestibles and livelihoods for most of the population in India. Considering past few years there has been strenuous effort to change the traditional agricultural practices and upgrade the practices in use. With the help of new technologies like artificial intelligence especially in the field of machine learning, by using these tools in the agricultural system to upgrade it. By considering the significant impact of weather, environmental changes, rainfall variability, water management, moister content, humidity in the soil and fertilizer application on crop, accurate fertilizer recommendations and crop output forecasting are imperative for agricultural shareholder. However, due to uncertain nature of these factors poses challenges in accurately predicting and analysing crop yield, especially amidst changing climatic conditions, often resulting in suboptimal productivity. In response to these challenges, researchers are increasingly turning to machine learning techniques to enhance crop productivity and sustainable growth.

This paper proposes a communal system designed to forecast crop yield and recommend fertilizers in the agricultural field. By harnessing agricultural datasets and employing algorithms such as the Extra Trees Regressor for yield prediction and Gaussian Naïve Bayes for fertilizer recommendation, our proposed approach aims to effectively enhance agricultural output.

II. LITERATURE SURVEY

The author P.M. Durairaj explore the challenges of yield prediction in agriculture. They leverage open data repositories and government agricultural databases to address these challenges. Their approach integrates deep reinforcement learning (DRL) and recurrent neural networks (RNN), particularly the Deep Recurrent Q-Network (DRQN). The proposed DRQN model achieves an impressive accuracy of approximately 93.7%. Furthermore, the paper indicates that the DRQN model outperforms existing models, showing promise for enhancing yield prediction accuracy in agriculture.[1]

The author Rajat Bishnoi and others have collected the dataset from the indiawaterportal.org website. They have used fuzzy logic to predict the yield of the crop as it operates on continuous range of value instead of discrete values, to process and rebuild the rainfall dataset. To predict the temperature and rainfall they have applied the auto regressive models. Mean by doing the square method along with calculating the error, the accuracy of the system is measured. [2]

In the article the authors rely on data is collected from indiawaterportal.org. Their primary objective is to classify soil samples based on their macro-nutrient content and predict the suitable crops for cultivation in that soil. To achieve this, they employ an ensemble classifier, which combines multiple distinct classifiers to provide improved predictive results compared to a single classifier model. The internal classifiers used include Naive Bayes Classifier, Random Forest Classifier, and CHAID Classifier, each suggesting appropriate crops, and considering soil conductivity, offering recommendations for suitable fertilizers to enhance soil fertility. [3]

The study addresses the pivotal issue of prediction of crop yield and suggestion of suitable fertilizer in the context of India's agriculture. Utilizing machine learning, particularly the Support Vector Machine (SVM) algorithm, the research aims to predict crop yields efficiently while offering precise fertilizer recommendations. The study focuses on important crops in the region, with data gathered from various sources over five years to support their predictive model. This research has the potential to significantly enhance agricultural

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY215

productivity and assist farmers in selecting suitable crops, ultimately contributing to food security in India. [4]

In their paper, the author address the challenge of forecast of crop yield and to suggest the fertilizer. With the help of Support Vector Machine and Random Forest to analyse five years of data related to agricultural, aiming to assist farmers in crop selection and fertilizer optimization. The Support Vector Machine achieves a yield prediction accuracy of approximately 99.47%, while the Random Forest algorithm achieves around 97.48%. [5]

III. OBJECTIVES AND GOALS

Create and put into place a system for decision-making process which uses machine learning practices to accurately gives us the amount of agricultural produce (crop yield), taking into account factors such as geographical location, changes that occur regularly during different seasons, crop types, and cultivation area.

To get the fertilizer recommendations for farmers based on single crop requirements and soil factors, targeting to improve resource usage for desired outcomes and reduce harm to the natural surroundings.

Improve agricultural output by using the predictive capabilities of the system to upgrade resource allotment, improve decision-making practices, and by the end to increase crop output.

Taking leading role within the agricultural sphere by combining different data-driven insights to lessen and reduce the ecological impacts and to improve and promote the optimal use of resources, and encourage the growth and development of more sustainable farming ecosystem.

IV. METHODOLOGY

The proposed agricultural system consolidates the machine learning concepts, mainly the Extra Trees Regressor for finding the final yield of the crop and Gaussian Naïve Bayes algorithm for demonstrate the most suitable fertilizer. The system intents to prophecy the output of the crop and endorse a fertilizer based on the feed given by the end user. Data that is selected for training the model is taken from source, which undergoes rigorous preprocessing to refine the raw content, and address the null vales. The model provides flexibility to the end users to choose between the crop yield and fertilizer recommendation system. The details like season, crop, area, region of the country, rainfall are considered to forecast the yield. Factors like humidity, soil nature, nitrogen, phosphorus and potassium levels, moisture elements in the soil are taken into consideration for endorse the satisfactory fertilizer.

For crop yield prediction, the Extra Trees Regressor is employed, by utilizing the dataset that is cleaned, used to train and test the model. User provided input data is taken into consideration to forecasts the final outcome based on the training given to the model by starting the process and providing the perfect outcome.

For suggesting the suitable fertilized to be used by the farmer, Gaussian Naïve Bayes algorithm is chosen because of its comprehensibility and capability. The model undergoes tailored preprocessing of data and for endorse the suitable fertilizer.

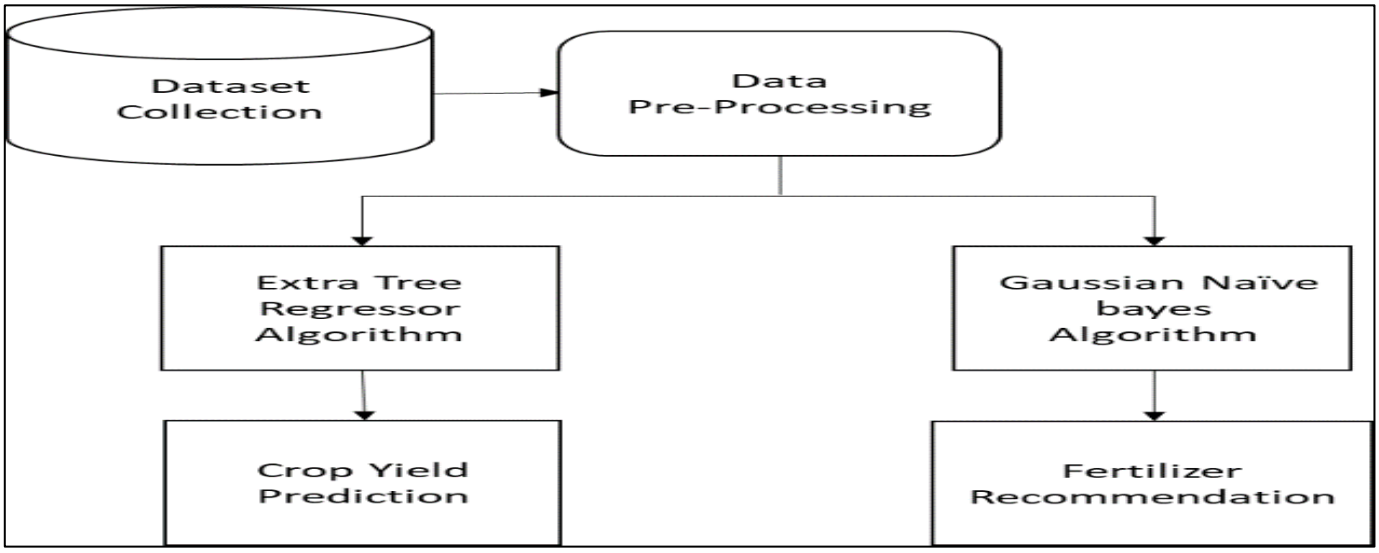

Fig 1: System Model

A. Datasets Description

We have made use of lots of data across India which counts up to two decades and provides a robust foundation to the system's functionality.

With the help of this dataset, for crop yield prediction the system uses six factors and to recommend the suitable fertilizer it considers six key factors. The inclusivity of the dataset enhances the efficiency and adaptability of the system.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY215

The dataset covers a larger group of crops and includes the most frequently grown crop across the country which is used for the further process. To make the preprocessing step expeditious, the crop yield and fertilizer data is converted into mathematical data to ensure the equilibrium of the system.

B. Extra Tree Regressor

Extra Trees Regressor is an extension of the Random Forest algorithm. Extra trees regressor is used for both regression and classification tasks. It boosts the process and is a part of ensemble model family as it typically consists of decision trees. In the first step the pre-processed dataset is given to the Extra tree regressor model as an input. the first iteration takes place. After the first iteration the dataset is classified as strong and weak learners. The weak learners are enhanced and added back to the strong learners which results in the updated dataset and this process is continued till it reaches the minimum error rates. As the weak learners are enhanced in the process, it provides the best efficiency for the system. Extra Trees Regressor works using fewer parameters and also reduces the risk of overfitting, which proves to a valuable addition to predictive model, especially in scenarios where accuracy is paramount.

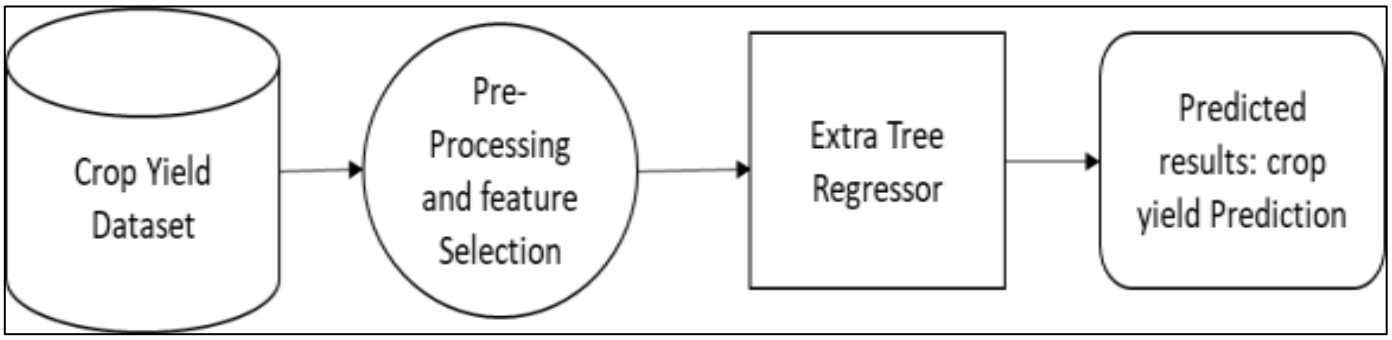

Fig 2: Extra Tree Flowchart

C. Gaussian Naive Bayes

Gaussian Naive Bayes, or GaussianNB, is a known supervised machine learning approach for various tasks, including regression and classification. By combining several decision trees, the ensemble approach GaussianNB improves performance of the prediction. The two main steps in the prediction process are: Average and voting on the various clustered decision trees. Though classification tasks depend on the majority vote for accurate prediction, the regression tasks use an average of the various model outputs in order to generate the reliable and accurate output. This algorithm uses the bagging concept in which it creates various subsets for different predictions by randomly selecting a replacement dataset from the original dataset for each decision tree from the original dataset. A reliable and accurate outcome is guaranteed when results are combined via majority vote. Its performance in classification tasks is remarkable which explains as to why it was chosen for this particular task. GaussianNB was chosen because it is ensemble-based approach, helps to decrease the variance and overcome the overfitting issues therefore increasing the accuracy. Also, this algorithm shows stability because any modifications to one decision tree will not substantially affect the others, and it is capable of handling the missing values on its own. Because of the resistance to noise, GaussianNB is a suitable choice for predictive modelling, particularly where precision and flexibility are needed.

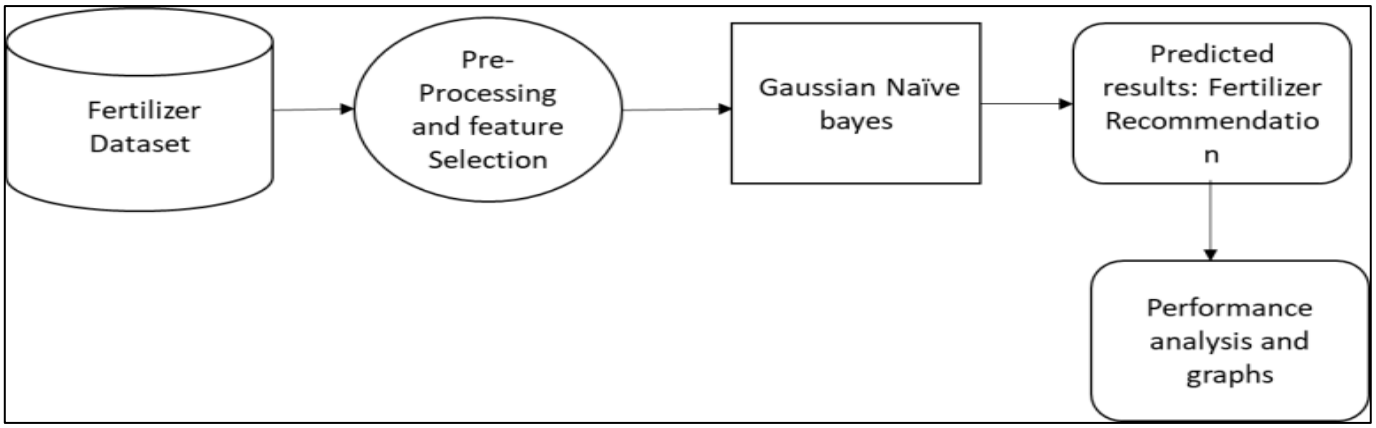

Fig 3: Gaussian NB Flowchart

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY215

V. RESULT AND ANALYSIS

A. Crop Yield Prediction:

```
#SAMPLE OUTPUT
#State_Name, District_Name, Season, Crop, Area
predictionss=etr.predict([[1,1,1,1254.0]])
A= predictionss[0]/1000
print("Yield : {0} tonnes".format(A))
```

```
Yield : 2.0305 tonnes
```

Fig 4: Prediction of Crop Yield – Output

```
print("Accuracy:%s"%(etr.score(X,y)*100))
```

```
Accuracy:99.88666042834282
```

Fig 5: Accuracy: Yield of The Crop

With the help of the datasets arbitrary values are passed as a first step of feed to the model.

B. Fertilizer Recommendation:

```
X1= [[26,52,38,4,3,37,0,0]]
p=model.predict(X1)
print("Fertilizer Name: " ,p[0])
```

```
Fertilizer Name:  Urea
```

Fig 6: Output: Recommending Suitable Fertilizer

Table 1: Accuracy of the Fertilizer System

	Accuracy	***TPR***	***F-Score***	***Support***
10-26-26	1	1	1	2
14-35-14	1	1	1	2
17-17-17	1	1	0.96	2
20-20	0.95	1	1	3
28-28	1	1	1	4
DAP	1	0.93	1	2
Urea	1	1	1	5
Weighted Mean	1	1	1	20
Macro Mean	1	1	1	20
Accuracy	1			

Once the yield is predicted, the next step is the recommendation of the suitable fertilizer. The model's performance inspection discloses a high level of accuracy, with an overall accuracy rate of 100%. The classification report lends a hand to demonstrate the model's effectiveness, with macro average precision, recall, and F1-score metrics all exceeding 90%. Specifically, the model achieves a macro average precision of 100%, a macro average recall of 100%, and a macro average F1-score of 100%. Moreover, the weighted average precision, recall, and F1-score metrics are also impressive, standing at 100%, 100%, and 100%, respectively.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY215

VI. CONCLUSION

The combined system, by making use of machine learning conceptualisation along with algorithms uses the Extra Trees Regressor and the Gaussian Naïve Bayes algorithms to forecast the crop yield and to endorse the suitable fertilizer, which provides better precision and performance. The scheme includes to build a comprehensible web application to set the seal on the ubiquitous accessibility for users. In addition to the Extra Trees Regressor model, the Gradient Boost algorithm is deliberation for intensifying the conjecture model. In place of Gaussian Naïve Bayes algorithm, Support Vector Machine and Decision Tree algorithms can be traversed to further refine the prediction model.

REFERENCES

[1]. Elavarasan, D., & Vincent, P. D. (2020). Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications. *IEEE access*, *8*, 86886-86901.

[2]. Bang, S., Bishnoi, R., Chauhan, A. S., Dixit, A. K., & Chawla, I. (2019, August). Fuzzy Logic based Crop Yield Prediction using Temperature and Rainfall parameters predicted through ARMA, SARIMA, and ARMAX models. In *2019 Twelfth international conference on contemporary computing (IC3)* (pp. 1-6). IEEE.

[3]. Archana, K., & Saranya, K. G. (2020). Crop yield prediction, forecasting and fertilizer recommendation using Data mining algorithm. *International Journal of Computer Science Engineering (IJCSE)*, *9*(1), 76-79.

[4]. Somwanshi, K., Sonawane, P. R., Lohar, T. S., & Jadhav, M. S. Crop Prediction and Fertilizer Recommendation Using Machine Learning.

[5]. Bondre, D. A., & Mahagaonkar, S. (2019). Prediction of crop yield and fertilizer recommendation using machine learning algorithms. *International Journal of Engineering Applied Sciences and Technology*, *4*(5), 371-376.

[6]. Zhang, X., Xu, M., Sun, N., Xiong, W., Huang, S., & Wu, L. (2016). Modelling and predicting crop yield, soil carbon and nitrogen stocks under climate change scenarios with fertiliser management in the North China Plain. *Geoderma*, *265*, 176-186.

[7]. Ghadge, R., Kulkarni, J., More, P., Nene, S., & Priya, R. L. (2018). Prediction of crop yield using machine learning. *Int. Res. J. Eng. Technol.(IRJET)*, *5*, 2237-2239.

[8]. [8] Filippi, P., Jones, E. J., Wimalathunge, N. S., Somarathna, P. D., Pozza, L. E., Ugbaje, S. U., ... & Bishop, T. F. (2019). An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. *Precision Agriculture*, *20*, 1015-1029.

[9]. Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and electronics in agriculture*, *151*, 61-69.

[10]. Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., ... & Kim, S. H. (2016). Random forests for global and regional crop yield predictions. *PloS one*, *11*(6), e0156571.