ISSN No:-2456-2165

# A Novel Strategy for Predicting Agriculture Crop and its Yield

Shobhit Tembhre[1] (Assistant Professor); Gopireddy Venkata Rukhmanand Reddy[2]; Chadive DevendraReddy[3]
Sudalakuntaakhil[4]; Penekalapati Pavan Kalyan[5]
School of Computer Science and Engineering, Presidency University, Bengaluru,India

**Abstract:- Agriculture holds a pivotal role in India, making substantial contributions to the nation's GDP (23%) and offering significant employment opportunities (59%), thereby securing the country's food needs. However, the sector faces challenges due to climate change, impacting crop predictability and resulting in lower yields. Recognizing the potential of technology to address these issues, particularly machine learning, this research project aims to guide inexperienced farmers in adopting advanced crop prediction methods. Machine literacy, a form of machine learning, is proposed as a solution to enhance agricultural practices by leveraging data and experience to make informed decisions, ultimately improving yields and sustainabilityin the face of climate change.**

## I. INTRODUCTION

Agriculture has traditionally been recognized as the primary source of daily necessities for humans, playing a pivotal role in India as a major industry and primary occupation. The promotion of ecological diversity and the production of nutritious crops has been advocated, urging farmers to adopt natural observation methods and abstain from using chemicals on both their fields and animals. Nevertheless, the swiftly changing modern climate presents a formidable obstacle, jeopardizing essential resources and undermining food supply and security.

Adding to these difficulties is the diminishing contribution of the agricultural sector to the nation's GDP. The sector's GDP share stood at approximately 17.2% in 2005, dwindled to 11.1 % in 2012, further decreased to 5% in 2018, and reached a meager 2% in the initial quarter of 2020. As 80% of farmers reside in rural areas, reductions in crop production revenue at the industry level significantly impact their way of life. The ongoing fluctuations in the agricultural sector and the resulting economic repercussions underscore the pressing need for sustainable practices and strategies to address the challenges posed by a changing climate.

In the context of crop prediction, machine learning leverages historical data from previous years and combines it with current monthly data to showcase the precision of meteorological information. Operating on the idea that computers can analyze data and makedecisions with minimal human input, machine learning serves as a method for automating models. A practical analogy is a logical classifier within machine learning, resembling a knowledgeable mathematician predicting the likelihood of a particular class's membership, such as estimating the probability of specific crop yields under givenconditions.

## II. LITEREACTURE REVIEW

- To achieve accurate harvest predictions, they propose the utilization of a supervised machine learning method employing a boosted naïve Bayes Gaussian classifier. This model predicts the output seed based on specific input parameters. Our initiative aims to greatly benefit less-experienced farmers lacking crop forecasting expertise for sustainable agriculture. Furthermore, future enhancements may involve providing tailored recommendations for fertilizers and guidelines suitable for different farms and crops, considering the input provided.[1]
- The utilization of a real-time dataset significantly improves the program's ability to make precise predictions, providing farmers with more valuable guidance for crop selection. This model, which incorporates artificial neural networks (ANN) and linear regression with both forward and backward propagation, demonstrated an 82% accuracy in predicting the dependent variable with minimal loss. The proposed backpropagation model, integrating gradient descent and the RELU activation function, is aimed at minimizing the mean square error (MSE). [2]
- The article delves into multiple machine learning techniques for forecasting agricultural production, focusing on variables like temperature, precipitation, season, and region. Through experiments utilizing datasets from the Indian government, it was determined that the Random Forest Regressor stands out as the most accurate predictor for crop yield forecasts. Furthermore, the study highlights that the Simple Recurrent Neural Network, functioning as a sequential model, demonstrates superior accuracy in predicting rainfall compared to Long Short-Term Memory (LSTM) models. [3]
- A model for forecasting soil series and suggesting an appropriate crop yield for that particular soil is put forth. Six upazila’s in the Khulna region's soil datasets were used for theresearch. The model has been put to the test using several machine learning algorithms. While KNN and bagged trees have decent accuracy, SVM has the greatest accuracy of all the classifiers when it comes to soil categorization. [4]

- The suggested approach facilitates the precise definition of features influencing crop yield. Its efficacy lies in combining a self- experimental analysis using Deep Q-Networks (DQN) with feature processing through Recurrent Neural Networks (RNN). Employing a Deep Recurrent Q-Network (DRQN) presents a comprehensive solution by autonomously uncovering the intricate relationships between crop yield and various environmental factors such as climate, soil, and groundwater. This contrasts with conventional supervised learning-based methods for crop yield prediction. The advantage of this method is a reduced relianceon expert input and historical knowledge during the model creation process for estimating crop productivity. Consequently, the proposed technique appears to employ a more comprehensive model for predicting crop yield.[5]
- Leveraging a real-time dataset improves the program's accuracy in making predictions, providing farmers with more valuable insights for crop selection. The model achieved an 82% accuracy in predicting the dependent variable with minimal loss, utilizing artificial neural networks (ANN) and linear regression incorporating forward and backward propagation. Additionally, the model includes a feature that assesses the crop success rate based on the farmer's input, ultimately suggesting the crop with the highestprobability of success. [6]
- The investigation leverages readily available data from government repositories, encompassing details such as state, district, crop, season, year, and production information. Various regression algorithms, including Random Forest, Gradient Boost, and Decision Tree, are tested, and ensemble methods are employed to enhance accuracy. Distinguishing itself from other studies, this research relies on easily accessible information such as state, district, and season, rendering it practical for farmers. The model demonstrates proficient accuracy in predicting crop yields, with noticeable improvement observed when utilizingensemble methods.[7]
- The study has successfully developed and deployed an intelligent crop recommendation system, accessible to farmers across India. This technology aids farmers in choosing the optimal crop for cultivation, taking into account diverse geographical and environmental parameters. Potential future enhancements may involve integrating a feature into the described model that anticipates crop rotations. Consequently, the system would consider the crop harvested in the previous cycle when recommending the current crop, ensuring maximum production.[8]

## III. METHODOLOGY

As commonly known, our application comprises two integral components: Machine Learning and Web Development. I will commence by delving into the foundational aspect of our application, which is the machine learning methodology. Before delving into thedetails, allow me to provide a concise overview.

➢ *Proposed System Design:*

The diagram provided below illustrates the fundamental configuration of the system we are currently designing (Figure 1). The initialstep involves the loading of an agriculture- related dataset. To improve accuracy, data preprocessing is undertaken, encompassing tasks such as filling in missing values and converting strings to numeric formats. The construction of the model incorporates machine learning techniques like Naive Bayes, Random Forest Regressor, and Linear

Regression. Following this, the model undergoes training using the pre-processed data. The accuracy of the model is subsequently evaluated through testing, employing various machine learning algorithms.

After determining the algorithm with the highest accuracy, the cultivation of crops willbe guided by the yield predictions generated by the model. It is important to note that our comprehension of the project is currently incomplete at this initial architectural phase, but these gaps are expected to diminish as theproject progresses.

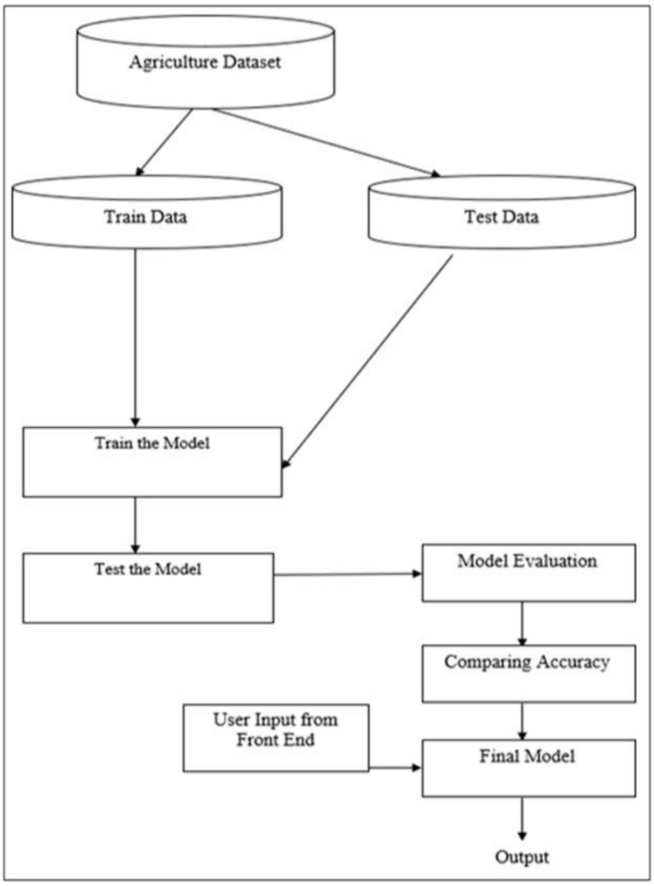


Fig 1 Proposed System Design Model

➢ *Data Flow Diagram:*

The diagram below, denoted as the Data Flow Diagram (Figure 2), illustrates the flow of data within the project. Farmers, acting as clients, initially input details like state name, district, season, area, and production into the user interface. These input values are then transmitted to the machine learning model. Utilizing the knowledge gained during training, the model takes on the responsibility of identifying the most suitable crop for achieving the highest yield. Furthermore, in cases where a lower yield is anticipated, the model suggests alternative crops for maximizing productivity. The user interface displays either the recommended crop or the projected yield based on the model's output.

For a more comprehensive grasp of the details provided, please consult the image presented below, indicated as Figure 2.

In our quest to connect the machine learning model with the webpage, we chose the Flask framework, a Python-based solution. The initial step involves converting the model into a pickle file for easy invocation and utilization. Subsequently, Flask is linked to the database, enabling the storage of client details. accuracy among various models. Opting for Random Forest as our final model was decided upon due to its superior accuracy.

The finalized model is now saved as a pickle file, prepared for future use when new data on crop or yield predictions becomes available. Following this, using HTML and CSS, we create a user interface or webpage. Users, upon visiting the webpage, can register for the first time by providing their details. This information is stored in the database, allowing for subsequent logins and contributions. The model utilizes these input values for predictions, and the results are showcased on the user interface. Integration of the database and our ultimate model with the user interface is facilitated through a Python Flask framework.

## IV. RESULT

- *Evaluation:*

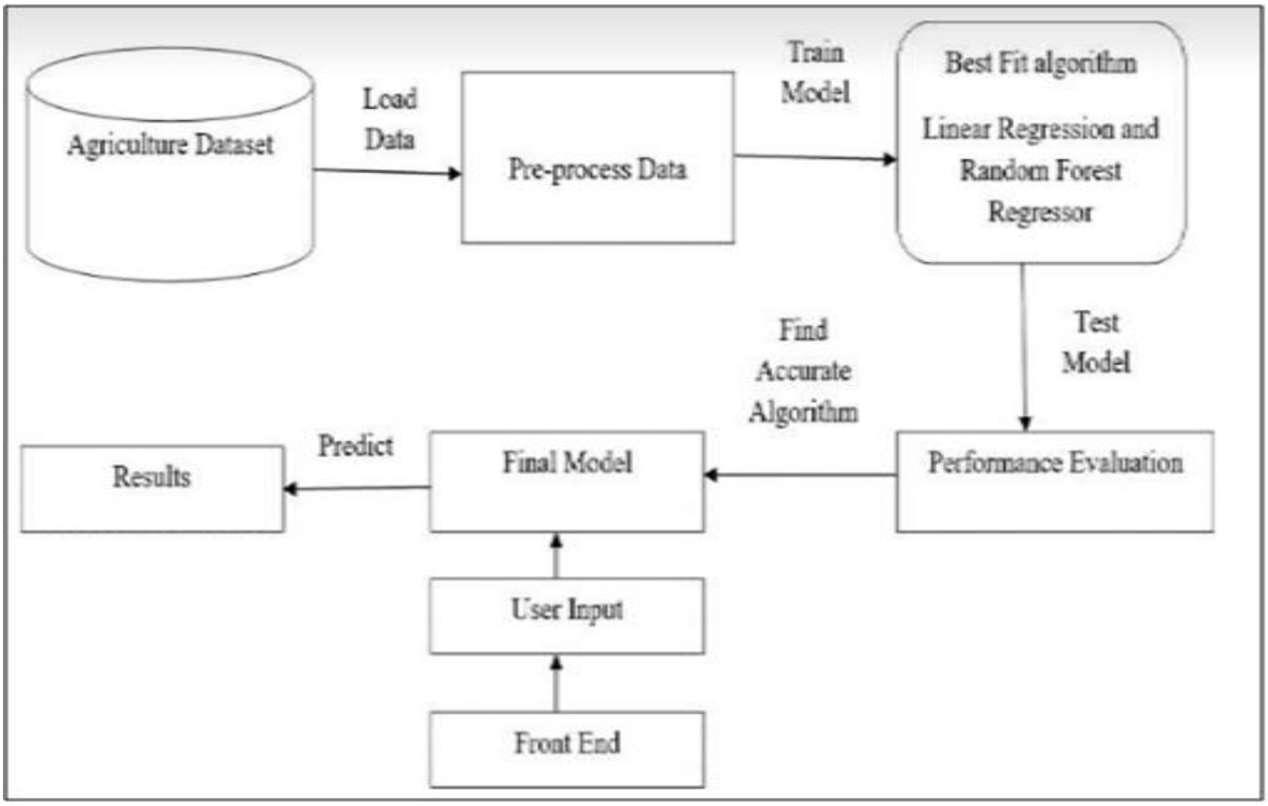


Fig 2 Data Flow Diagram

Let's initiate our examination by focusing on the backend. We utilized a range of machine learning techniques, including XGB Regressor, Random Forest, Decision Tree Classifier, and Linear Regression, to craft our model. After preprocessing the data, we created train and test datasets. The model underwent training using thetrain dataset, employing the mentioned algorithms to establish multiple models. Python served as the programming language in this process. Evaluating the model's accuracy, we selected the final model based on the highest accuracy, ensuring more precise predictions for both crop type and yield. Diverse techniques were applied to enhance the model's accuracy, and Figure (3) illustrates a comparison of In the testing phase, the model effectively predicts the class label, providing insights into the recommended crop and the expected yield. Ongoing evaluations of the model's performance involve the utilization of test data, with a subsequent comparison of the results. The testing process has yielded notable success, achieving an accuracy rate of 98% when applying the model to the test dataset.

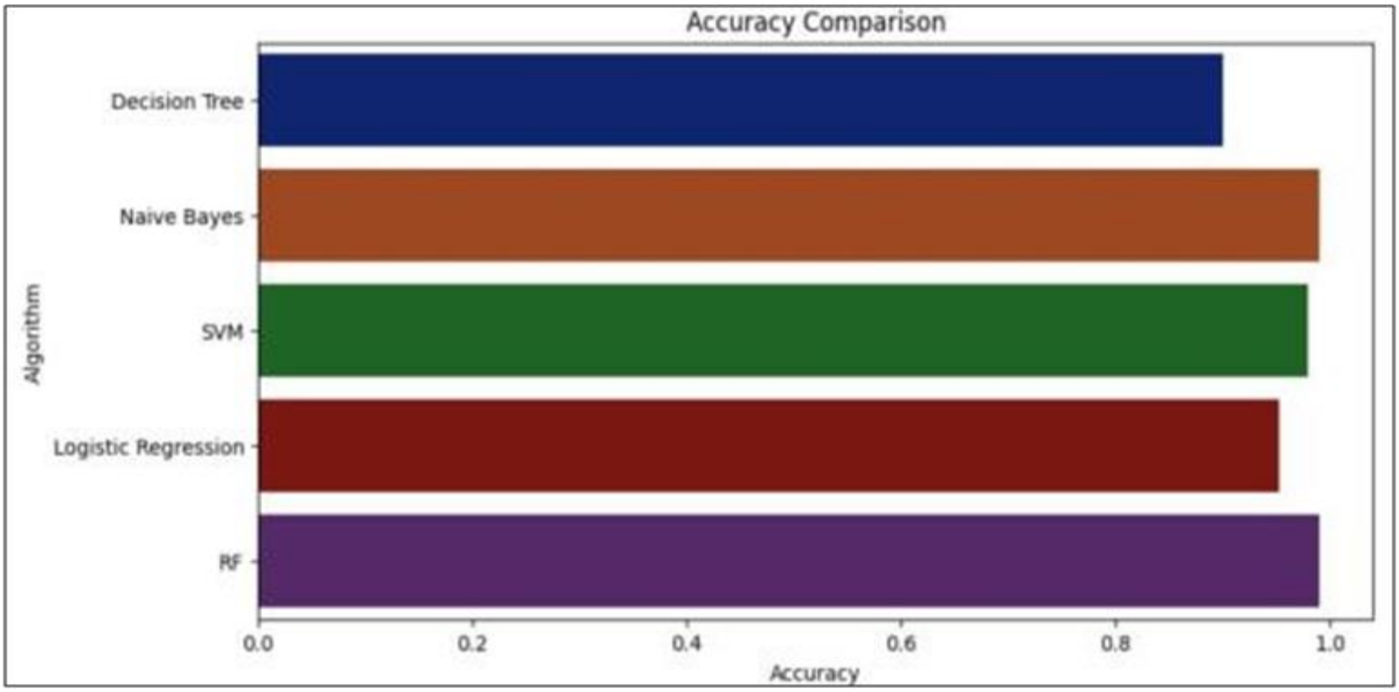

Fig 3 Accuracy Comparison

## V. CONCLUSION

The agricultural sector is facing significant challenges related to global sustainability, which calls for creative approaches to improve productivity while protecting environmental resources. The project showcases the efficacy of machine learning in providing tailored crop recommendations and yield prediction based on different parameters. Its success signifies a step towards optimizing agricultural practices, promoting sustainability, and offering valuabledecision-making tools for farmers in cultivatingcrops best suited for their specific conditions. Continued advancements in this field hold significant potential for revolutionizing farmingpractices and addressing agricultural challenges globally. The primary driving forces behind crop prediction and agriculture initiatives are the need to maintain food security in the face of an expanding global population, maximize resource utilization, and boost productivity. We are happy with our 98%accuracy rate.

## FUTURE SCOPE

Further developments in artificial intelligence and machine learning algorithms may result in increasingly complex crop prediction models. Better algorithms can shed more light on the variables affecting crop yield, growth, and disease resistance. Incorporating blockchain technology into the agricultural supply chain can improve traceability and transparency. This makes sure the information used to forecast crops is reliable and contributes to thedevelopment of a more dependable system. Working with emerging AgTech companies cancreate new opportunities for implementing cutting-edge tools and technologies into agricultural and crop forecasting initiatives. Startups frequently offer original concepts and cutting-edge methods for solving problems.

## REFERENCES

[1]. M. Kalimuthu, P. Vaishnavi and M. Kishore, "Crop Prediction using Machine Learning," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp.926-932, doi:10.1109/ICSSIT 48917.2020.9214190.

[2]. R. Medar, V. S. Rajpurohit and S. Shweta, "Crop Yield Prediction using Machine Learning Techniques," 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India, 2019, pp. 1-5, doi: 10.1109/I2CT45611.2019.9033611.

[3]. A. Nigam, S. Garg, A. Agrawal and P. Agrawal, "Crop Yield Prediction Using Machine Learning Algorithms," 2019 Fifth International Conference on Image Information Processing (ICIIP), Shimla, India, 2019, pp. 125-130,doi: 10.1109/ICIIP47207.2019.8985951.

[4]. S. A. Z. Rahman, K. Chandra Mitra and S. M. Mohidul Islam, "Soil Classification Using Machine Learning Methods and Crop Suggestion Based on Soil Series," 2018 21st International Conference of Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2018, pp. 1-4, doi: 10.1109/ICCITECHN.2018.863194

[5]. D. Elavarasan and P. M. D. Vincent, "Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications," in IEEE Access, vol. 8, pp. 86886-86901,2020, doi: 10.1109/ACCESS.2020.2992480.

[6]. S. S. Kale and P. S. Patil, "A Machine Learning Approach to Predict Crop Yield and Success Rate," 2019 IEEE Pune Section International Conference (PuneCon), Pune, India, 2019, pp. 1-5, doi: 10.1109/PuneCon46936.2019.9105741.

[7]. P. S. Nishant, P. Sai Venkat, B. L. Avinash and B. Jabber, "Crop Yield Prediction based on Indian Agriculture using Machine Learning," 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 2020, pp. 1-4, doi: 10.1109/INCET49848.2020.9154036.

[8]. Z. Doshi, S. Nadkarni, R. Agrawal and N. Shah, "AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1-6, doi: 10.1109/ICCUBEA.2018.8697349.