International Journal of Innovative Research in Engineering
Volume 3, Issue 6 (November-December 2022), PP: 71-74.
www.theijire.com **ISSN No: 2582-8746**

Crop Yield Prediction based on Indian Agriculture Using Machine Learning

Govind Yadav[1], Chandrabhan Yadav[2], Ashish Maurya[3], Shailesh Patel[4]
[1,2,3] Student, Department of Computer Science and Engineering, ITM Gida Gorakhpur, India.
[4] Assistant Professor, Department of Computer Science and Engineering, ITM Gida Gorakhpur, India.

How to cite this paper:
Govind Yadav[1], Chandrabhan Yadav[2], Ashish Maurya[3], Shailesh Patel[4], "Crop Yield Prediction based on Indian Agriculture Using Machine Learning", IJIRE-V3I06-71-74.

Abstract: *Agriculture, as everyone in India knows, is the bedrock of our country. The yield of almost all crops farmed or planted in India is predicted in this article. In this script, the user can predict the yield of the crop in the year that they want to produce or plant by using simple factors such as State, District, Season, and Area. The paper predicts the yield using advanced regression techniques such as the Random Forest Algorithm (RFA) of Supervised Learning and the concept of Regression to improve the algorithms. Technology advancements must be worked on across multiple disciplines, and it has already shown improvements in many fields. It will display the nitrogen, potassium, and phosphorus content ratio in the soil.*

Key Word: *Random Forest Algorithm, Farming, Machine Learning Techniques, Proper Yield, Indian Agriculture.*

I. INTRODUCTION

In our research, which we found in the previous research papers is that everyone using the climatic factors like soil type, rainfall, sunlight nutrients possessed by the soil (Nitrogen, Phosphorus etc.) but the problem scenario is that we need to gather the data and then a third party does this prediction and then it is explained to the farmer and this takes a lot of time and effort for the farmer and he doesn't understand the science behind these technology. This study employs straightforward variables, such as the farmer's state and district of residence, the crop being grown, and the season, in order to make it straightforward and directly applicable to farmers (as in Kharif, Rabi, etc.) .In India, there are more than the hundred of crops are planted around our whole country. Crop is categorized betters understanding & visualizations. This data of the research have been acquire from the Indian Government Repository [1]. With almost 4.5 Lakhs observations, the data includes many parameters including State, District, Crop, Season, Year, Area, and Production .The many Indian states and territories that show which crop category is well-known throughout which season. To reduce error and produce precise predictions of various crops, we applied advanced regression approaches including the Random Forest Algorithm (RFA), which is based on Supervised Learning .India's primary and most significant community .Since ancient people grow the crops on their own agricultural land, and they are able to fulfill their needs. As a result, natural crops are grown and used by a number of species, including humans, livestock, and birds. The creature has taken the greenish goods created in the soil, which has result in a healthy and welfare life. The agricultural sector has been steadily degrading since the advent of new advanced technologies and techniques.

II. LITERATURE SURVEY

Crop Yield prediction studies have been conducted using various machine learning techniques and algorithms that can be used by agriculture institutions. This paper gives an overview of some of these studies conducted in research papers using the techniques and results they used.

1. Ananthara, M. G. et al. on (2013, February) "Survey On Crop yield Prediction On Indian Agriculture Using Machine Learning" ,[1] It proposed a prediction model for datasets pertaining to agriculture which is called as SVM algorithm for crop yield using regenerative techniques. They considered parameters likely crop type, soil & soil pH value, humidity and crop sensitivity. Their analysis was mainly in wet, rice and the sugarcane are yields in India. The proposed algorithm was Support Vector Machines SVM algorithm and it will be performed well with the accuracy of 85 percent.

2. Awan, A. M. et al. (2006, April) "Survey On Crop yield Prediction On Indian Agriculture Using Machine Learning" ,[2] Built an new and smart framework focused on crop yielding prediction clustering decision tree methodology and then considered parameters like plantation, temperature and precipitations of the rainfall in their different regions. The experimented weighted of decision tree methodology with the spatial constraint for the analysis of palms field.

3. Chawla, I. et al. (2019, August) "Survey On Crop yield Prediction On Indian Agriculture Using Machine Learning"[3] Used the fuzzy logic for crop yield prediction through the statistical time series models. They considered the different parameters

like rainfall & temperature for the predictions. The prediction was classified with levels 'good yield', 'very good yield', & 'poor yield'.

4. Chaudhari, A. N. et al. (2018, August) "Online Survey On Crop yield Prediction on Agriculture Using the Machine Learning",[4] Used two algorithms namely clustering k-means and Bias algorithm, then they uses the many others algorithm for the better utilization of crop yield prediction and they also considered the parameters like different Area region, Rainfall, Soil type Ph value, and also that the system was able to tells that which crop is very suitable for cultivation based on the mention features.

5. Manjula, A et al. "Weather Crop yield Prediction On Agriculture land Using Machine Learning", [5] Built a crop selection and predict the yield which will considered the various index like vegetation, temperature humidity and difference vegetation factors, Then distinguish between the climate factor and agro factors and other disturbances caused by the prediction for the better understanding .

6. Tripathy, A. K. et al. (2016, July) "Survey On Crop yield Prediction On Indian Agriculture Using Machine Learning", [6] Were some of the authors uses the Support Vector Machines (SVM) algorithm to predict the high crop yielding with same features as in the previous paper mention, applied for SVM and neural networks for high crop yield prediction proposed a new decision system which is an interface to give the input and get the better output. They also considered the climatic factors mainly (considering range) temperature, precipitation and reference crop evapotranspiration.

7. Ishwar Chawla et al. employing the ARMA, SARIMA, and ARMAX models to predict temperature and rainfall factors, fuzzy logic-based crop yield prediction: [7] In India's economy, agriculture is quite significant. Forecasting crop yields is so crucial to India's prosperity. Numerous weather factors, such as temperature and rainfall, have an impact on crops. As a result, it's crucial to consider these aspects when estimating a crop's yield. It's challenging to predict the weather. In this study, ARMA, SARIMA, and ARMAX (Auto Regressive Integrated Moving Average) are used as three forecasting techniques (ARMA with exogenous variables). The output of the three models is compared, and the most accurate model is utilized to predict temperature and rainfall, which are subsequently utilized to anticipate crop yield.

III.METHODOLOGY

The algorithm used in our project are Random Forest Algorithm (RFA), and Neural networks. The measured results of our experiments show that the RFA algorithm outperforms other algorithm, and that it reaches the highest accuracy not only in weather prediction also help in the high yielding of crops.

According to the soil quality, the system estimates the crop output and offers fertilizer recommendations when necessary. The user must input the site and pH value (calculated based on the % of nutrition). Two controllers handle the processing of the results. Address is used as an input to controller, in addition to the use of third-party applications such APIs for weather and temperature, type of soil, nutrient value of the soil in the region, amount of rainfall in the region, and soil composition are frequently calculated. Then it receives a pH value as input, and uses that value to determine the soil's alkalinity. Additionally, the proportion of nutrients such as nitrogen (N), phosphorus (P) & potassium (K).

Neural Network

Neural networks, a subset of machine learning that are often referred to as artificial neural networks (ANNs) or simulated neural networks (SNNs), are the foundation of deep learning approaches. Their organization and terminology are modeled after the human brain, mimicking how real neurons communicate. An input layer, one or more hidden layers, and an output layer make up a node layer in an artificial neural network (ANN). Each node, or artificial neuron, is interconnected with others and comes with a weight and threshold. Any node whose output rises above the specified threshold value is activated and starts sending information to the top layer of the network.

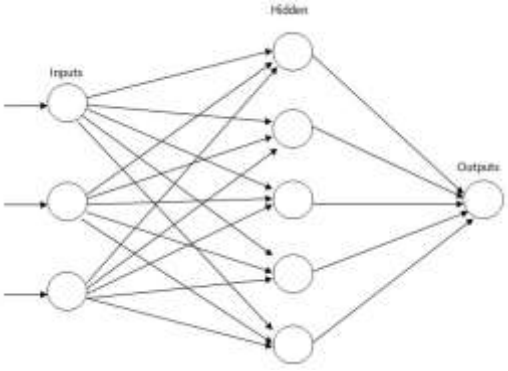

Figure-1 [Neural Network]

Random Forest Algorithm

A well-known machine learning algorithm Random Forest is a technique used in the supervised learning methodology. It can be used to solve ML problems involving both classification and regression. It is based on the concept of ensemble learning, which is a method of combining different classifiers to address difficult issues and improve model performance. Random Forest is a classifier that, as the name implies, uses a number of decision trees on different subsets of the provided dataset and averages them to improve the dataset's predictive accuracy. Instead of relying on a single decision tree, the random forest uses forecasts from each tree and predicts the outcome based on the majority of predictions' votes.

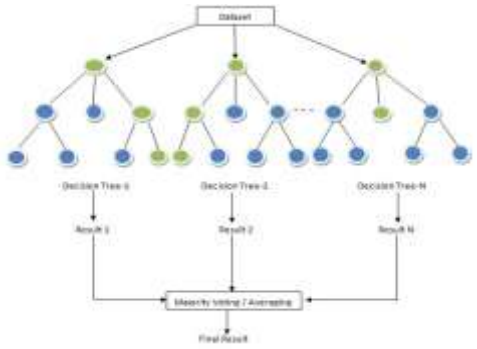

Figure-2 [Random Forest Algorithm]

Existing System & Disadvantages

There is no computerized system to identify crop recommendations or behavior analysis in the current system. In the first place, it is only appropriate for instance-level techniques that call for an instance classifier, As we previously discussed, popular current neural network methods employ separated instances as inputs and then transform them into embedding space using a deep neural network.

Advanced System

The suggested method A machine learning model is used to estimate the ideal crop variety for the chosen plot of land because environmental variables vary by region. Using data from the standard dataset to train the crop suggestion model, machine learning techniques are utilized to select the best crop to cultivate with the highest possibility of growing. Utilizing Naive Bias, and Support Random forest machine techniques, the best crop type is determined. Based on this model, it was decided what kinds of crops the farmer should grow. Aspects to take into account include humidity, temperature, soil moisture, and Rainfall level.

Advantages

1- System that is efficient and secure.

2- The benefit of this approach for agriculture is that it allows us to choose the right crops based on factors like climate, temperature, etc.

3- Machine learning avoids assumptions, the challenges of employing larger sample sizes, and complex problems by assisting us in making predictions using the available data.

IV.CONCLUSION

This project uses machine learning and employs the methods of Random Forest, Polynomial Regression, and Decision Tree to assess performance. Random forest provides the best yield prediction among the three algorithms in our suggested model when compared to other methods. The output that indicates improvements in the dataset is classified using the Random Forest Regression Technique, and Neural Network models. So, after analysis, we determined that the proposed model for determining crop yield is more effective than the current approach. The establishment of the aforementioned method would aid in improving our nation's agriculture practices. Additionally, it can be applied to lessen farmer losses and boost crop yields to increase agricultural capital. By integrating, the model can be made better. When we use stacked regression, the outcome is more unpredictable than when we use each model alone.

References

[1]. Ananthara, M. G., Hemavathy, R., & Arunkumar, T. (2013, February). CRY is a more accurate model for predicting crop yields from agricultural data sets utilizing bee hive clustering .In 2013 International Conference on Informatics, Mobile Engineering and Pattern Recognition.

[2]. Sap, M. N. M., and Awan, A. M. (2006, April). a kernel-based intelligent system for forecasting crop yields, Conference on knowledge Discovery and Data Mining in Pacific-Asia.

[3]. S. Bang, R. Bishnoi, A. S. Chauhan, A. K. Dixit, & I. Chawla (2019, August). Crop Yield Prediction Based on Fuzzy Logic Using Predicted Temperature and Rainfall Parameters from ARMA, SARIMA, and ARMAX Models. The 12th International Conference on Contemporary Computing (IC3) will be held in 2019.

[4] R. Jaikla, S. Auephanwiriyakul, & A. Jintrawet (2008, May) .Using support vector regression, estimate rice yield. The Fifth International Conference on Electrical Engineering/Electronics, Computer, Telecommunication, and Information Technology was held in 2008.

[5] Miniappan, N., Kadir, M. K. A., and Ayob, M. Z. (August 2014). Prediction of wheat yield using artificial neural networks. The fourth International Conference on Engineering, Technology, and Entrepreneurship (ICE2T) was held in 2014.

[6]. KM Meheta (January 2015). XCYPF: A adaptable and scalable framework for crop yield prediction in agriculture. 9th International Conference on Intelligent Systems and Control (ISCO) of the IEEE was held in 2015.

[7]. Paul, M., S. K. Vishwakarma, and A. Verma (2015, December). Data mining approach for soil behavior analysis and crop yield prediction. 2015's Computational Intelligence and Communication Networks International Conference (CICN) (pp. 766-771). IEEE.

[8]. Shah A, Dubey A, Hemnani V, Gala D, and Kalbande D. R. (2018) .Crop Yield Prediction Using Regression Techniques in a Smart Farming System. International Conference on Wireless Communication Proceedings (pp. 49-56). Singapore's Springer.

[9]. Ramalatha, M., Suresh, A., and Kumar, P. G. (2018, October). Tamil Nadu's main agricultural yields are predicted using K-means and Modified KNN. The third International Conference on Communication and Electronics Systems (ICCES) was held in 2018. (pp. 88-93). IEEE.

[10]. Misra, B., Veenadhari, S., and Singh, C. C. (2014, January). A method using machine learning for predicting crop yield based on climate factors .Global Conference on Computer Communication and Informatics in 2014.