# INTERNATIONAL JOURNAL
# FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

# Agriculture Crop Yield Prediction Using Machine Learning

Firdous Hina[1], Dr. Mohd. Tahseenul Hasan[2]
[1, 2]Anjuman College of Engineering and Technology, Sadar, Nagpur-17

Abstract: In our suggested system, we employed a vast dataset that included all of India's states, whereas in the old system, just a single state was considered. These suggestions may be extracted and used to educate the farmers. The farmer can have a better understanding of the crops to cultivate by using a pictorial depiction. Machine Learning Techniques develops a well-defined model with the data and helps us to attain predictions. Agricultural issues like crop prediction, rotation, water requirement, fertilizer requirement and protection can be solved. Due to the variable climatic factors of the environment, there is a necessity to have a efficient technique to facilitate the crop cultivation and to lend a hand to the farmers in their production and management. This may help upcoming agriculturalists to have a better agriculture. A system of recommendations can be provided to a farmer to help them in crop cultivation with the help of data mining. To implement such an approach, crops are recommended based on its climatic factors and quantity. Data Analytics paves a way to evolve useful extraction from agricultural database. Crop Dataset has been analyzed and recommendation of crops is done based on productivity and season
Keywords: Machine learning, Agriculture techniques, crop predictions

## I. INTRODUCTION

Agricultural unpredictability, due to changing temperature season soil parameters are the factors which reduces productivity. More production should be gained with a larger population and area, however this is impossible to achieve. Farmers used to rely on word-of-mouth, but owing to climate reasons, they can no longer do so. Agricultural elements and characteristics are used to provide data that may be used to learn more about Agri-facts. AgricultureSciences presents key highlights in the IT industry to assist farmers with good agricultural information. In the current situation, the ability to apply new technical approaches to the sector of agriculture is desirable. MachineLearning Techniques use data to create a well-defined model that may be used to make predictions. Crop forecast, rotation, water requirements, fertilizer requirements, and crop protection are all challenges that may be resolved. All learning takes place in the brain of machine learning. The machine learns in the same way that humans do. Humans gain knowledge from their experiences. The more information we have, the easier it is to make predictions. By analogy, the chances of success are lower in an unknown circumstance than in a known situation. Machines are taught in the same way. The machine examines an example in order to produce an accurate forecast. When we give the machine a comparable scenario, it can predict what would happen. However, much like a person, if the computer is fed an example that has never been seen before, it has trouble predicting what will happen next.

Learning and inference are the two main goals of machine learning. First and foremost, the machine learns by recognising patterns. Machine Learning programmes have a clear life cycle that may be stated as follows:

1) Formulate a query
2) Gather information
3) Visualize data
4) Train the algorithm
5) Put the algorithm to the test
6) Gather feedback
7) Fine-tune the algorithm
8) Repeat steps 4–7 until the results are satisfactory.
9) Make a forecast using the model.

Once the algorithm has mastered generating the correct conclusions, it applies what it has learned to fresh data sets.. To collect and process data, technologies such as blockchain, IoT, machine learning, deep learning, cloud computing, and edge computing can be used. Computer vision, machine learning, and IoT applications will assist farmers and related domains enhance production, improve quality, and ultimately increase profitability.

Pre-harvesting, harvesting, and post-harvesting are the three major categories of agricultural activity. Machine learning advancements have aided in boosting agricultural gains. Machine learning is a recent technology that is assisting farmers in reducing farming losses by offering detailed crop recommendations and insights.. Machine learning has by far been the most helpful and an innovative method that is recommended to be used in agriculture based applications. A farmer can receive an accurate estimate of harvestable versus non-harvestable acres on a particular day by logging into a customised dashboard on a computer or tablet using machine learning technology. Harvestable crop weight and maturity can also be assessed and projected. Additionally, crops can be examined both before and after harvest for the presence of attractive traits, level of damage (if applicable), nutritional makeup, and other factors that may effect the ultimate viable yield and product pricing using a range of technologies, including image analysis. Machine learning (ML) approaches are used in a number of businesses, from anticipating consumer phone usage to analysing customer behaviour in supermarkets (Ayodele, 2010). (Witten et al., 2016). Machine learning has been used in agriculture for some years (McQueen et al., 1995). Crop production prediction is one of the most challenging challenges in precision agriculture, and various models have been proposed and verified thus far. This task involves the use of numerous datasets since crop production is impacted by a number of factors such as climate, weather, soil, fertiliser use, and seed type (Xu et al., 2019). This implies that estimating agricultural output is not a simple process, but rather a succession of complex steps. Crop yield prediction programmes can already estimate actual yields with reasonable accuracy, although better yield prediction performance is still sought. Researchers have helped to solve the problem of climate change's influence on agricultural output in a variety of ways. In this section, we provide a quick assessment of the work of colleagues and attempt to critically analyse it in order to define the area of future study. We will continue our investigation as part of our research .Its applications can be divided into four major categories:

*a)* Crop management
*b)* Livestock management
*c)* Water management
*d)* Soil management

## II.      RELATED WORKS

LIsmail et al. (2018) [1] created a system for predicting a country's crisis readiness. Climate change is being studied using machine learning. Southeast Asia is the focus of the study. The processes involved in calculating the predictive index include data collecting, data training, data testing, and index computation. An index can be used for prediction, index validation, and index visualisation, among other things. The study is being carried out as a prophylactic measure. The regions will be notified, and the susceptible index will be checked using deep learning.

## III.      PROPOSED WORK

Crop yield forecast is a critical challenge for decision-makers at all levels, including national and regional (e.g., EU) decision-making. Farmers can use an accurate crop production prediction model to assist them determine what to grow and when to grow it Crop yield forecasting is a difficult task for decision-makers at all levels, including at the national and regional (e.g., EU) levels. Farmers may use a crop production prediction model to help them decide what to plant and when to plant it. Crop production forecasting may be done in a number of different methods. This review paper looks at what has been written on the use of machine learning in agricultural yield prediction in the literature. The fact that the publication is a survey or typical review piece is one of the exclusion criteria in our analysis of the retrieved publications. Those articles that were left out are really part of a larger project discussed in this section.

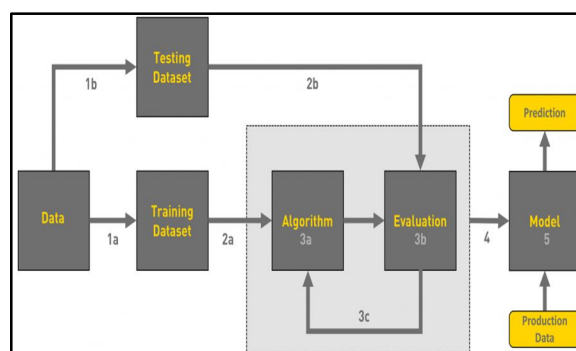## IV.      MACHINE LEARNING IMPLEMENTATION



Fig i Machine Learning block diagram

Python Libraries required for implemention
1) Numpy
2) Pandas
3) Sci-kit Learn
4) Matplotlib

A. *Pre-processing of Data*
The technique we choose to collect data is determined on the type of project we wish to build. For example, if we want to construct an ML project that leverages real-time data, we may use data from numerous sensors to create an IoT system. The data set can originate from a variety of sources, including a file, a database, a sensor, and many others, but it cannot be used for analysis right away because it may contain a large amount of missing data, exceptionally large values, disorganised text data, or noisy data. As a result, Data Preparation has been performed in order to address this problem. Data pre-processing is one of the most important steps in machine learning. It's the most important stage in making machine learning models more accurate. In machine learning, there is an 80/20 rule. Every data scientist should spend 80% of their time pre-processing data and 20% of their time on actual analysis.

## V. DATA CLASSIFICATION

When the goal variable is categorical (i.e., the output may be classified into classes – it belongs to Class A, B, or something else), a classification problem arises.
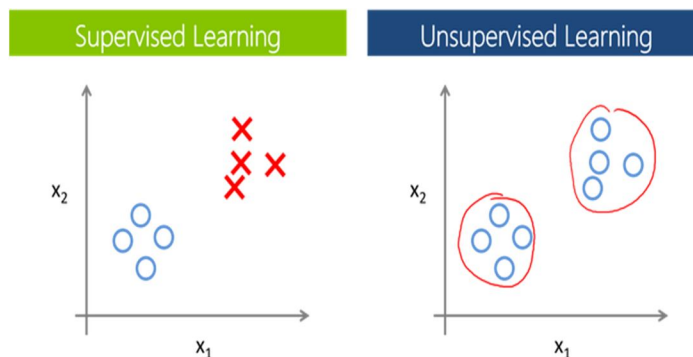


Fig ii: Classification techniques

## VI. ALGORITHMS FOR CLASSIFICATION

A. K-Nearest Neighbor
B. Naive Bayes
C. Decision Trees/Random Forest
D. Support Vector Machine
E. Logistic Regression

We split a model into three sections to start training it: 'Training data,' 'Validation data,' and 'Testing data.' You train the classifier with a 'training data set,' tweak the parameters with a 'validation set,' and evaluate the classifier's performance with a 'unseen test data set.' It's worth remembering that during training, the classifier only has access to the training and/or validation sets. When the classifier is being trained, the test data set must not be used. Only while the classifier is being evaluated will the test set be available..
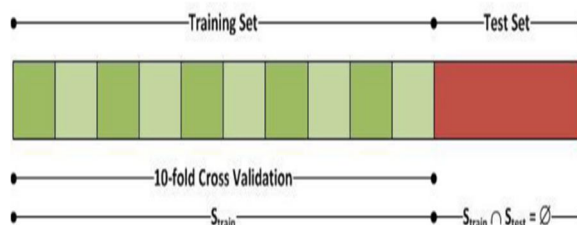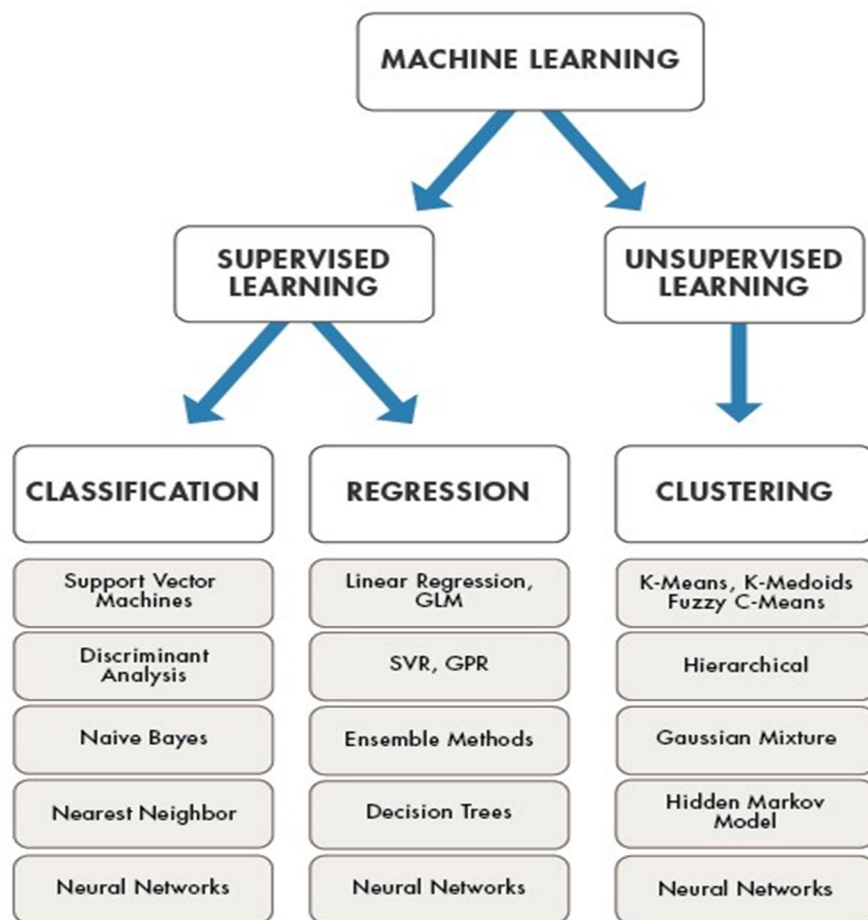
Fig iii: Algorithms for classification

The project's problem statement is to use a Decision Tree Classifier to propose crops to farmers. The main method of this project is to preprocess the data supplied to us, then utilise it to construct the model for the backend and link it to the UI interface using flask to display the entire and final product. In our suggested system, we employed a vast dataset that included all of India's states, whereas in the old system, just a single state was considered. These suggestions may be extracted and used to educate the farmers. The farmer can have a better understanding of the crops to cultivate by using a pictorial depiction. There are 821 unique data points in the dataset. The dataset has 14 columns, which are explained below.

1) States: India's total number of states
2) Ground Water: Total ground water level
3) Rainfall: rainfall in millimetres
4) Temperature: the temperature is measured in degrees Celsius.
5) Soil type: There are a variety of soil kinds.
6) Season: When is the best time to plant crops?
7) Crops: Different types of crops
8) Fertilizers needed: Types of Fertilizers Needed
9) Cultivation cost: Total Cultivation Cost
10) Revenues forecasted: Total revenue forecasted
11) Seeds per hectare amount: seeds per hectare quantity
12) Culture duration: amount of days for cultivation length
13) Crop demand (nineteenth): crop demand (ninth): crop demand (ninth
14) Crops for mixed cropping: which crops can be cropped together
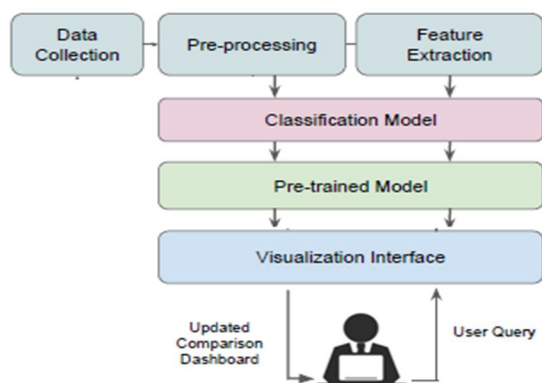
Fig iv: Algorithms for classification

Gather information and prepare it for training. Clean anything that may require it (remove duplicates, correct errors, deal with missing values, normalization, data type conversions, etc.) Randomize data to remove the impacts of the sequence in which we gathered and/or otherwise prepared our data. Perform further exploratory analysis, such as visualising data to assist uncover meaningful associations between variables or class imbalances (bias alert. Sets for training and evaluation have been created.
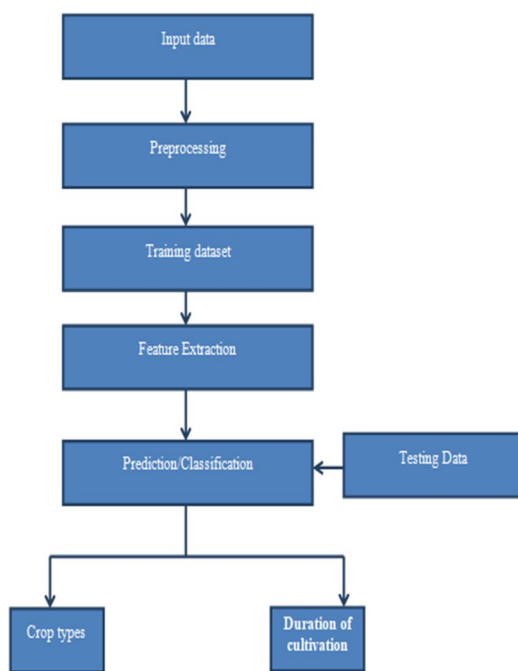


Fig v: Flowchart

## VII. CONCLUSION

The importance of crop management was extensively addressed in this research. Farmers require modern technologies to help them raise their crops. Agriculturists can be notified about accurate crop predictions on a timely basis. The agricultural factors were analyzed using a variety of Machine Learning approaches. The features chosen are determined by the dataset's availability and the research's goal. According to studies, models with more characteristics may not always deliver the highest yield prediction performance. Models with more and fewer features should be evaluated to discover the best performing model. Several algorithms have been employed in various research. The findings reveal that while no definitive conclusion can be taken about which model is the best, they do show that some machine learning models are utilized more frequently than others. With the incorporated model we found an average accuracy of 95%.

## REFERENCES

[1] Shreya S. Bhanose, Kalyani A. Bogawar (2016) "Crop And YieldPrediction Model", International Journal of Advance ScientificResearch and Engineering Trends, Volume 1,Issue 1, April 2016

[2] A.T.M.S. Ahamed, N.T. Mahmood, N. Hossain, M.T. Kabir, K. Das, F. Rahman, R.M. Rahman Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh 2015 IEEE/ACIS

[3] I. Ahmad, U. Saeed, M. Fahad, A. Ullah, , A. Ahmad, J. Judge **Yield forecasting of spring maize using remote sensing and crop modeling** (10) (2018)

[4] I. Ali, F. Cawkwell, E. Dwyer, S. Green**Modeling managed grassland biomass estimation by using multitemporal remote sensing data—a machine learning approach**IEEE J. 10 (7) (2017),

*[5]* Cheng, H.; Damerow, L.; Sun, Y.; Blanke, M. Early Yield Prediction Using Image Analysis of Apple Fruit and Tree Canopy Features with Neural Networks. *J. Imaging* **2017**,

[6] Chen, Y.; Lee, W.S.; Gan, H.; Peres, N.; Fraisse, C.; Zhang, Y.; He, Y. Strawberry Yield Prediction Based on a Deep Neural Network Using High-Resolution Aerial Orthoimages. Remote Sens. 2019,

[7] Črtomir, R., Urška, C., Stanislav, T. et al. Application of Neural Networks and Image Visualization for Early Forecast of Apple Yield. Erwerbs-Obstbau 54, 69–76 (2012).

[8] Duo Attention with Deep Learning on Tomato Yield Prediction and Factor InterpretationPRICAI 2019: Trends in Artificial Intelligence, 2019,

[9] 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,

[10] Tripathy, A. K., et al.(2011) "Data mining and wireless sensornetwork for agriculture pest/disease predictions." Information andCommunication Technologies (WICT), 2011 World Congress on.IEEE.

[11] PriyankaP.Chandak (2017)," Smart Farming System Using DataMining", International Journal of Applied Engineering Research,Volume 12,

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)