**Introduction.**

The main challenge faced in the agriculture sector is the lack of knowledge about the changing variations in climate.Each crop has its own suitable climatic features. This can be handled with the help of precise farming techniques. The precision farming not only maintains the productivity of crops but also increases the yield rate of production. In India it is very important to maintain sustainable agriculture to meet growing needs. Though many steps have been carried out to minimize the loss of crops, traditional methods has its own disadvantages.

These disadvantages can be overcome with the help of precision farming. With the use of IOT and prediction system, precision farming makes decision. The data from the field area are collected using IOT system. The prediction framework is fed with the collected data from the sensors for obtaining suggestion. Crop selection and changing climatic conditions are the two major problems faced among the farmers. With the existing prediction and monitoring methods we can address this problem. Though these methods are useful, there is no optimum solution for the crop suggestion. Some of the drawbacks found in the existing system are the improper analysis, choosing effective algorithms, and efficient selection of attributes all these parameters may affect the crop yield. The proposed system helps in overcoming the drawbacks found in the existing system. The methods in the proposed system includes increasing the yield of crops, real-time analysis of crops using IOT, selecting efficient parameters, making smarter decisions and getting better yield. Effective algorithms need to be used for early prediction of crops. By choosing correct factors most useful data can be generated using ML model by intake of various parameters and this generated data helps the farmers by suggesting a right crop to be sown. Suggesting the precise crop for a particular field area is the main aim of crop suggestion model. By choosing the suitable crops for the field area we can minimize the loss of crops. Suitable algorithms with selective features has to be chosen since the accuracy level of suggesting crops may vary based on the type of algorithm chosen.

Digital Farming and Precision Agriculture allow precise utilization of inputs like seed, water, pesticides, and fertilizers at the right time to the crop for maximizing productivity, quality and yields. Most of the farmers practice traditional farming patterns to decide crops to be cultivated in a field. Thus, the solution to this Challenge can be a crop recommendation system that helps farmers to decide the right crop to sow in their field and forecast the yield & revenue.

**Novelty.**

We have implemented a high end machine algorithm for the precise determination of the type of crops that can be grown over an area and its projected yield. With the aid of IBM WATSON our proposed solution also gives personalized suggestions to the farmers regarding the type of fertilizers that can be used and the temperature requirements of each type of crop. This helps farmers to ease the production work so that they get a better yield.

**Smart Irrigation.**

To cope with the shortage of water, there is a critical need for some smart irrigation systems that can irrigate more areas with low consumption of water. A comprehensive review of various techniques of smart irrigation has been presented in Jha et al.However, there is an availability of various low water consumption-based irrigation techniques, for example, sprinkler systems and drip irrigation systems; but these systems need human intervention up to a great extent. There is a scope to add features to existing systems to develop smart irrigation systems. The system continuously monitors the level of water in a crop, compares the water content available in soil and crop plant with standard need of water. It automatically starts sprinklers or drips as per water requirements of the crop.

**Crop Disease Prediction and Health Monitoring.**

AI-based techniques with inclusion of image processing, deep learning, and data analysis provide an easy and effective way for disease prediction and health monitoring. The system captures crop images using high definition(HD) camera-enabled drones, unmanned aerial vehicles (UAVs), or satellite imagery. The captured images are used as a dataset for training of Convolutional Neural Networks (CNNs), a class of Artificial Neural Network (ANN). CNNs extract useful features from fed images and make predictions about disease(s) in crops. The system is effective in continuous monitoring of the health of plants and hence gives better solutions for calculating the amount of pesticides to be used and time to use a pesticide.

**Crop Readiness Identification.**

AI-based system captures images of a crop and analyzes them for determining the crop readiness in a particular area for harvesting. The crops can be categorized into different categories on the basis of readiness and other quality parameters before actually sending them to market. For categorization, use of various pattern clustering techniques viz. K-means, fuzzy C-means (FCM), expectation maximization (EM), and hierarchical clustering plays an important role.

**Yield Prediction.**

Yield prediction is an area of interest for researchers for past many decades. Yield prediction requires yield mapping devices, which are still not easily available to farmers. CNNs can solve the problem of yield prediction in an economic and easy way. Many researchers [10–12] have developed models to predict crop yields using AI-based methods which use RGB/normalized difference vegetation index

(NDVI) images. Their experimental results on publicly available dataset prove the usefulness of their models. A comparative analysis of CNN-based models with AI- based models clearly indicates that CNNs are more advantageous than traditional Machine Learning (ML) and AI-based techniques [13]. Authors will discuss the detailed process of predicting crop yields using CNN.

**Weed/Pest Management**

Traditionally, weed management was accomplished with a combination of many techniques viz. mechanical weed control, crop rotation, herbicides, etc. In addition to the above, various biochemical-based combined weed and pest control solutions are also available in the market, but they all result in reducing the yield productivity. The reduction in crop yield occurs due to the spraying of pesticides frequently and in a constant amount in a field. AI provides an intelligent solution for this challenge.

**Precision Farming/Agriculture**

The aim of precision agriculture (PA) is to increase yield production as well as quality by simultaneously reducing the overall cost and environmental pollution . The quantity and quality of a crop depend on many parameters such as soil, weather, irrigation, etc. So, there is a requirement to monitor all these parameters at a regular interval of time. Traditional monitoring techniques are not adequate in accurate and efficient monitoring of these parameters. So, there is vital demand for an automated system which can perform monitoring of parameters in an effective way. In the modern era, proximity sensing and remote sensing dominate the field of agriculture and effectively monitor a plethora of parameters required for better prediction and planning of agricultural practices. Proximity sensing specifically deals with soil using high-resolution data. Remote sensing provides the geographical sensing of fields using various sensors. Thermal remote sensing provides some additional information such as temperature, water status, etc. It also helps us in getting many vegetation indices.

**Business and Social impact**

Let’s take a look at the benefits of crop yield prediction with AI/ML

* Remote monitoring of farm areas
* Resource and field mapping
* Crop monitoring for better yield
* Predictive analytics for making data-driven decisions
* Weather forecast .

Businesses from the AgTech industry today are making use of neural network algorithms to predict crop yield. The backpropagation algorithms help in identifying the appropriate weight value of the yield to calculate the error derivative. Accuracy of crop yield estimation is significant for agronomic production reasons.

So, predicting crop yield is essential for the food production ecosystem around the world. With better data in hand, it becomes possible to make informed decisions. Government agencies also find the crop yield prediction data useful as they can plan accordingly for national food security.

1. Traditional agriculture cycle

1. Precision agriculture cycle.

**Tech Stack**

Weather prediction

Open weather API

The Regression models used for this project:

* Gradient Boosting Regressor
* Random Forest Regressor
* SVM
* Decision Tree Regressor.

**IBM Technologies used.**

IBM Watson

IBM Cloud.

(Categorization of machine learning algorithms.)

Prediction of soil properties is the first and the most crucial step which influences the selection of crop, land preparation, selection of seed, crop yield, and selection of fertilizers/manure. The soil properties are directly related to the geographic and climatic conditions of the land in use and hence is an important factor to take into consideration. The soil properties prediction mostly consists of predicting nutrients in the soil, soil surface humidity, weather conditions during the lifecycle of the crop.

                                         b) Proposed System Architecture.

**1. RANDOM TREE:**

Random tree is similar to that of a decision tree.But it differs from random tree in a way that for each split only a random subset of attributes are available. Random tree can be built for both nominal and

numerical data. The Random tree is similar to C4.5 or CART butt it selects only a random subset of

attributes. At each node it considers K randomly chosen attributes. The subset ratio parameter

specifies the size of the subset.

**2. K-NEAREST NEIGHBOR:**

K-Nearest Neighbour can be used for both classification and regression. K-Nearest Neighbours is a non-complex algorithm which stores all the available cases and classifiers new cased based on

some similarity measure. The sampled set is classified based upon the ”closeness” sssthat is the distance measure such as Euclidean distance or Manhattan distance.

**3. RANDOM FOREST:**

Random forest square measure associate ensemble learning methodology for classification, regression and different tasks, that operate by building a mess of call trees at coaching time and outputting the category that’s the mode of categories or mean prediction of the individual trees. Random call forests correct for call tree custom of over fitting to their coaching set. The primary rule for random call forest was created by Tin KamHo victimization the random mathematical space methodology, which, in Ho’s formulation, could be a thanks to implement the “stochastic discrimination” approach to classification.

**4. DECISION TREE:**

Classifies data using the attributes. Tree consists of decision nodes and decision leafs. Nodes can have two or more branches which represents the value for the attributes tested. Leafs nodes produces a homogeneous result.

**Classification**

**Sum Pooling Layer**

Like Max and Average pooling layer, sum pooling layer is used to calculate the sum all the values from each of the group of input data and the output is the sum of all the values from each group of the input.

**Architecture of a general yield prediction system**

**Block diagram of Crop Suggestion System**

**IOT System Design**

**Scope of work**

Automated Soil and Crop Monitoring

LANDSAT data has helped in agricultural monitoring since 1972. Whether it is about estimating crop production or monitoring water usage, LANDSAT imagery is of great importance. Besides, they are also beneficial in field-level management to identify different conditions and increasing yield through zone mapping.

Weather Forecasting

Predicting the weather is possible through IoT-based sensors that use historical data. The past data will always have a pattern, based on which the system will predict future weather conditions. It becomes ideal for farmers as they can find a suitable time to sow seeds and harvest the crop.

**Crop Health Analysis**

Geospatial AI analysis help in monitoring and identifying crop health through drone-based imagery data. Drones capture field data and transmit it to computers for further analysis. The system has algorithms to analyse images and understand farm health. The benefit is the identification of pests so that mitigating measures help overcome the problem.

**Predictive Analytics**

It has become a vital aspect of precision farming in recent years. IoT devices can study crop rotation, water management, pest attacks, nutrition management, and much more. The devices then generate rich insights through spatial analysis that aid improved standards of crop production.

**Remote Sensing-Based Yield Prediction**

Remote sensing brings information about a field, crops, or an object without actually visiting the field. The information is captured with the help of different sensors and is made available in raw, processed, or analyzed form. It works on the principle of properties of objects. The properties may be chemical, physical, structural, energy emissions etc.

**Data-Driven Management for Advanced Farming: Principal Stages.**

The raw measurements of key parameters from crops need to be efficiently processed so that numbers or images unambiguously turn into valuable information. Crop management based on field data already evolved when Precision Agriculture came to light thirty years ago, but it has certainly been transformed by the present digital information era. Traditionally, and in those places where technology has not arrived yet, field management consists of visually inspecting the development of crops to reach a diagnosis with which farmers make decisions and actuate giving different treatments to their crops. This approach relies on field experience and the information perceived through the eyes of farmers.

**Grid maps of NDVI (Normalized Difference Vegetation Index) without zoning (a), and after applying a clustering algorithm.**

**Model Building Procedure**

To tackle this given problem we have used Regression Analysis.

Regression analysis is a form of predictive modelling technique which investigates the relationship between a dependent (target) and independent variable (s) (predictor). The Regression models used for this project:

* Gradient Boosting Regressor
* Random Forest Regressor
* SVM
* Decision Tree Regressor

**Step 1.**

Data Collection.

**Data Source:** <https://data.worldbank.org/indicator/AG.YLD.CREL.KG>

**Step 2**

Data Pre-processing and Data Analysis

**i) Loading the dataset**

**ii)dropping some of the columns such as Area Code, Domain, Item Code, etc, that won’t be of any use to the analysis. Also, renaming Value to hg/ha\_yield**

**iii) using describe () function in pandas we get an overall insight**

**iv)Data Exploration**

Grouping by countries and items to have a better understanding of the dataset. There are 101 countries in the dataframe, with India having the highest crop yield production. In addition, India is the highest for production of cassava and potatoes. Potatoes seem to be the dominant crop in the dataset, being the highest in 4 countries.

v) Using one hot encoding we convert categorical variables into the form ML can process and predict.

vi) Data Visualization

Feature importance

The boxplot below shows the yield for each item. Potatoes are the highest, Cassava, sweet potatoes, and Yams.

**Step 3.**

**Model Creation.**

In this step we will create 4 different ML models **u**sing DT, SVM, Gradient Boosting, Random forest Algorithm.

**Step 4**

**Comparing and selecting the model**.

The evaluation metric is set based on R² (coefficient of determination)

**Conclusion**

From the above result we have concluded that Decision Tree is the best model with R² value of 96%. India is a nation in which agriculture plays a prime role. In the prosperity of the farmers, the nation prospers. Thus our work would help farmers in sowing the right seed based on soil requirements to increase productivity of the nation. Our future work is aimed at an improved data set with a large number of attributes and also implements yield prediction.By utilizing both ANN and CNN algorithms for harvest yield it helps to ensure the precision and effectiveness of crop yield for maximizing crop production.