**ML Crop Prediction**

From ancient period, agriculture is considered as the main and the foremost culture practiced in India. Ancient people cultivate the crops in their own land and so they have been accommodated to their needs. Since the invention of new innovative technologies and techniques in the agriculture field is slowly degrading. Due to these, abundant invention people are been concentrated on cultivating artificial products that is hybrid products where there leads to an unhealthy life. Nowadays, modern people don’t have awareness about the cultivation of the crops in a right time and at a right place. Because of these cultivating techniques the seasonal climatic conditions are also being changed against the fundamental assets like soil, water and air which lead to insecurity of food.

The machine learning learns the algorithm based on the supervised, unsupervised, and Reinforcement learning each has their importance and limitations. Supervised learningthe algorithm builds a mathematical model from a set of data that contains both the inputs and the desired outputs. Unsupervised learning-the algorithm builds a mathematical model from a set of data which contains only inputs and no desired output labels. Semi-supervised learning- algorithms develop mathematical models from incomplete training data, where a portion of the sample input doesn't have labels.

**Need of Work :**

In the past, crop prediction was performed by considering farmer's experience on particular field and crop. The crop prediction is a major issue that remains to be solved based on available data. Machine learning techniques are the better choice for this purpose. Different Machine learning techniques are used and evaluated in agriculture for estimating the future year's crop production.

Our system implements a system to recommend crop from previous data. This is achieved by applyingRandom Forest machine learning algorithmon agriculture data and recommends fertilizer suitable for every particular crop. Our system focuses on creation of a prediction model which may be used for future prediction of crop. It presents a brief analysis of crop prediction using machine learning techniques.

**PROPOSED SYSTEM:**

In the past, crop prediction was performed by considering farmer's experience on particular field and crop. The crop prediction is a major issue that remains to be solved based on available data. Machine learning techniques are the better choice for this purpose. Different Machine learning techniques are used and evaluated in agriculture for estimating the future year's crop production.

Our system implements a system to recommend crop from previous data. This is achieved by applying Random Forest machine learning algorithm on agriculture data and recommends fertilizer suitable for every particular crop. Our system focuses on creation of a prediction model which may be used for future prediction of crop. It presents a brief analysis of crop prediction using machine learning techniques.

The aim of proposed system is to help farmers to cultivate crop for better yield. The crops selected in this work are based on important crops from selected location. The selected crops are Rice, Jowar, Wheat, Soybeans, and Sunflower, Cotton, Sugarcane, Tobacco, Onion, Dry Chili etc. The dataset of crop yield is collected from last 5 years from different sources.

Crop Prediction can be done using crop data, nutrients and location data. These inputs are passed to Random Forest and Support Vector Machine algorithms. These algorithms will predict crop based on present inputs.

**What is Random Forest Regression?**

Every decision tree has high variance, but when we combine all of them together in parallel then the resultant variance is low as each decision tree gets perfectly trained on that particular sample data, and hence the output doesn’t depend on one decision tree but on multiple decision trees. In the case of a classification problem, the final output is taken by using the majority voting classifier. In the case of a regression problem, the final output is the mean of all the outputs. This part is called Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap and Aggregation, commonly known as **bagging**. The basic idea behind this is to combine multiple decision trees in determining the final output rather than relying on individual decision trees.   
Random Forest has multiple decision trees as base learning models. We randomly perform row sampling and feature sampling from the dataset forming sample datasets for every model. This part is called Bootstrap.

We need to approach the Random Forest regression technique like any other machine learning technique

* Design a specific question or data and get the source to determine the required data.
* Make sure the data is in an accessible format else convert it to the required format.
* Specify all noticeable anomalies and missing data points that may be required to achieve the required data.
* Create a machine learning model
* Set the baseline model that you want to achieve
* Train the data machine learning model.
* Provide an insight into the model with test data
* Now compare the performance metrics of both the test data and the predicted data from the model.
* If it doesn’t satisfy your expectations, you can try improving your model accordingly or dating your data, or using another data modeling technique.
* At this stage, you interpret the data you have gained and report accordingly.

***OutPut***





