Forecasting and Examination of Crop life using Artificial Intelligence and Machine Learning Surface Energy Balance Algorithm

[1]N. Sreeja, [2]K. Helini Reddy, [3]Dr.P.Vinay Bhushan, [4]S.Sandeep Babu

[1]Assistant Professor, Department of CSE (DS), Vignan's Institute of Management and Technology for Women, Kondapur, Ghatkesar, Telangana

[2]Assistant Professor, Department of Information Technology, Vignan's Institute of Management and Technology for Women, Kondapur, Ghatkesar, Telangana

[3]Associate Professor, Department of CSE, School of Engineering, Malla Reddy University, Hyderabad

[4]Assistant Professor, Department of CSE (AI&ML), Vignan's Institute of Management and Technology for Women, Kondapur, Ghatkesar, Telangana

Abstract

This Paper relates to the research area of crop yield prediction, and it provides better decision making in farm management and planning. Pesticide's quantity and dosages are not being considered in the existing studies. Based on studies, the proposed work is focused on prediction of crop pesticides requirement based on ground conditions and its impact on plant cultivation. So, it is necessary to consider the dosages and it gives better information for different crops along with pesticides dosages and this Paper proposes a model and compute reduction of pesticide dosages by introducing the compost pit calculation and tells best crop yield based on season and area and analyses the moisture content for each crop using Artificial Intelligence, Machine Learning, Surface Energy Balance Algorithms.

Keywords: *Crop Prediction, Artificial Intelligence, Machine Learning, Surface Energy Balance Algorithm.*

1. Introduction

Agriculture is one of the main supporting sectors of the Indian economy and most of the rural population depends on it for livelihood. India is a country that is rich in terms of food and environmental resources. Nevertheless, such prosperity is gradually reducing and resulting inflow agricultural productivity and low income for the farmer. Farmers additionally lacked insights into agricultural strategies and high quality production planning. The

agriculture sector provides about half the amount i. e. , 52 percent of the total number of jobs available in India, approximately. It is the primary means of living for almost two-thirds of the population in India. From the statistics given by the economic data of the financial year 2019-20, agriculture has acquired 24 percent of India's GDP. The agriculture sector occupies around 43 percent of India's geographical area. Flood, drought, and hailstorms are some of the natural calamities that affect agriculture production, causing threats to the livelihoods of farmers and their food security. However, the bitter fact to digest is that around 10-12 farmers end up their lives every month by committing suicide due to failure of crops, bore wells, and debt burdens. Weather condition plays a vital role in agricultural production. It has a great influence on the growth, development, and yields of a crop. Crop yield prediction is the biggest farming problem. Every cultivator must know how much harvest he is about to expect. In the olden days, to achieve better crop yield, the farmers were aware of the earlier familiarity with an exacting crop. In an Indian agriculture context, there is a huge volume of data that should be maintained, and it could be turned into information which is extremely valuable for a lot of purposes. The methodology proposed here helps the farmers to predict the crop yield and profit in prior. Various parameters that affect crop growth were analysed in detail. Using the results given for a crop, a farmer can be aware of their net profit before sowing the crop.

2. Related Work

This Paper relates to the research area of crop yield prediction, and it provides better decision making in farm management and planning. Some similar research areas are, forecasting of agricultural yields in the Indus Basin, Pakistan by Bastian’s and Ali by combining three models, such as surface energy balance algorithm for land (SEBAL), Carnegie Institution Stanford model, and Monteith's model. Monteith’s model is used to compute absorbed photosynthetically active radiation (APAR). For the characterization of land wetness dynamics and Spatio-temporal variability, the Carnegie Institution Stanford model is used. SEBAL is used to find the efficiency of light usage. This combined model takes inputs and crop identification maps as generic meteorological measurements [1][2]. To calculate the values, GIS environment and satellite measurements were used. In order to achieve crop yield prediction, NOAA-AVHRR measurements are used. For crop yield forecasting, Kastens et al. , used remote sensing images and

performed Image masking, which is nothing but limiting the study to a subset of pixels in the given area, which is their approach to achieve crop yield prediction. They considered imagery of 1 KM for experiments. They proved that yield correlation masking is a viable alternative to cropland masking with respect to crop yield prediction[3][4].

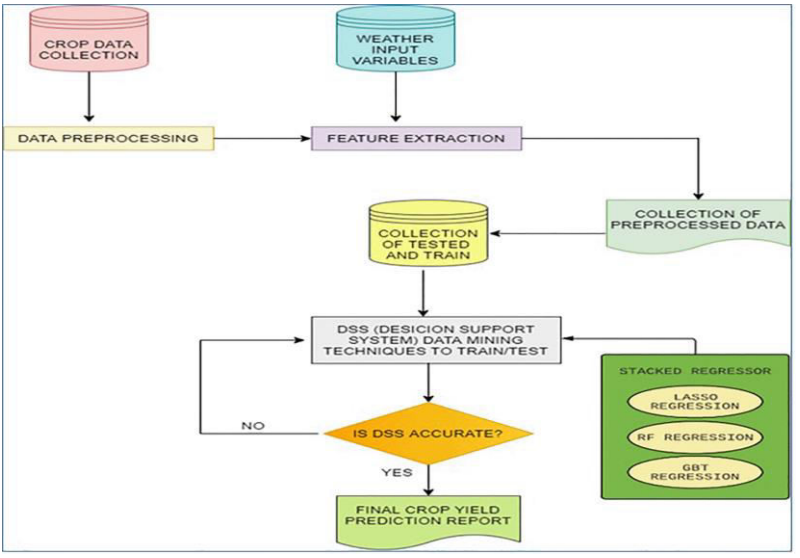

Figure1: Process of Crop Prediction

3. Methodology

This Analytics which will tell amount the harmful chemicals in pesticides which cause the death of crops. Consider the following data:

We are going to consider the crops and their chemical quantity for the states in India from January 2017 to May2020. These data include different type of chemicals relying in the pesticides and damaging the crop and its growth[5][6]. According to the Pollution board of Central Government, conditions specified on the crops based on their quantity and affect rate of the pollutants such as PM2. 5, PM10, NO, NO2, Nox, NH3, CO, SO2, O3, BENZENE, TOULENE, XYLENE.

Table-1: Major-pollutants (csv datafile)

	1	2	3	4	5
1	PM2.5-AVG	PM10-AVG	NO2-AVG	NH3-AVG	SO2-AG
2	190	131	107	4	42
3	188	131	110	4	40
4	280	174	155	2	37
5	302	181	144	2	39
6	285	160	121	3	19
7	252	191	143	2	33
8	273	157	76	5	17
9	126	111	49	6	14
10	127	112	35	5	17

Table-2: Crop pollutant states(CSV data)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		CITY	DATE	PM2.5	PM10	NO	NO2	Nox	NH3	CO	SO2	O3	Benzene	Toluene	Xylene	Crop Severity
2	1969	Amaravati	25-11-2017	81.4	124.5	1.44	20.5	12.08	10.72	0.12	15.24	127.09	0.2	6.5	0.06	Moderate
3	1970	Amaravati	26-11-2017	78.32	129.06	1.26	26	14.85	10.28	0.14	26.96	117.44	0.22	7.95	0.08	Moderate
4	1971	Amaravati	27-11-2017	88.76	135.32	6.6	30.85	21.77	12.91	0.11	33.59	111.81	0.29	7.63	0.12	Moderate
5	1972	Amaravati	28-11-2017	64.18	104.09	2.56	28.07	17.01	11.42	0.09	19	138.18	0.17	5.02	0.07	Moderate
6	1973	Amaravati	29-11-2017	72.47	114.84	5.23	23.2	16.59	12.25	0.16	10.55	109.74	0.21	4.71	0.08	Moderate
7	1974	Amaravati	30-11-2017	69.8	114.86	4.69	20.17	14.54	10.95	0.12	14.07	118.09	0.16	3.52	0.06	Moderate
8	1975	Amaravati	01-12-2017	73.96	113.56	4.58	19.29	13.97	10.95	0.1	13.9	123.8	0.17	2.85	0.04	Moderate
9	1976	Amaravati	02-12-2017	89.9	140.2	7.71	26.19	19.87	13.12	0.1	19.37	128.73	0.25	2.79	0.07	Moderate
10	1977	Amaravati	03-12-2017	87.14	130.52	0.97	21.31	12.12	14.36	0.15	11.41	114.8	0.23	3.82	0.04	Poor

In Predictive Analytics we will use sklearn module which is supported by Python. In the first phase, the data is collected from the Ponnur Rural Test Agriculture along with Crop ID with different set of Attributes. Then data is merged based on Crop ID which act as an Primary key, when coming to data Cleaning Process the 'Number_weeks_Used' column contains missing values. So that column missing values are replaced by mean of that column. Immediately we can move to data/feature extraction. Feature can also be called as an attribute. Extraction means deriving the data from the existing attribute. As the ID attribute is no longer helpful to analyze the data, we are going to delete the ID and this process is known as

Feature Deletion. We are selecting remaining features for prediction. Sowe can include the remaining features. This is known as Feature Selection. Finally, we can apply the Machine Learning algorithms[7][8].

In this work, we have carried predictive analytics on three datasets with the following attributes.

Dataset1:Estimated_Insects_Count, Crop_Type, Soil_Type, Pesticide_Use_Category, Number_Doses_Week, Number_Weeks_Used, Number_Weeks_Quit, Season, TotalDoses, Doses_Week_Used_Quit.

Dataset 2: Crop, Soil, Pesticide type, Quantity, Daysused.

Dataset 3: Year, State, Watershed_Code, Pesticidesrange.

We Obtained data set I and II from the following individual data by merging based on ID (Crop– id)

a. Crop_yielding

This dataset describes about Cost of Cultivation and yield across different states and comprises of 49 records and 5 attributes.

Table-3: Crop_yielding

Crop	Object
State	Object
Cost of Cultivation	Float
Cost of Production	Float
Yield	Float

b. Crop damage

This describes about crop damage based on Pesticide usage category and includes 88858 records and 5 columns.

Table-4: Crop damage

ID	Object
Estimated_Insects_count	Int
Pesticide_Use_Ctgry	Int

Season	Int
Crop_damage	Int

c. Crops data

This dataset describes about different types of crops based on ID. It comprises of 89958 records and 2 columns.

Table-5: Crops data

ID	Object
Crop	Object

d. Doses

This dataset describes the number of pesticides doses used per week, which includes 88858 records and 4 attributes.

Table-6: Doses

ID	Object
No_Doses_Week	Int
No_Weeks_Used	Float
No_Weeks_Quit	Int

e. Pesticides

This dataset includes chemical compound used over different states and quantity used.

Table-7: Pesticides

Compound	Object
State_Code	Int
Quantity	Float

f. Soil type

It includes 88858 records and 3 attributes which contain ID, Crop_type and Soil-type.

Table-8: Soil type

Compound	Object
State_Code	Int
Quantity	Float

4. Experiments and Results

It is saying that in data set I, crop damage is high and to reduce that hazardous chemical, our Paper-paper is going to suggest the usage of chemical pesticides in a low quantity level. For that we created a model to predict the crop life based on pesticides, crop details and soon.

Data set II tells the crop life based on the pesticides used with respect to the days. But the major difference between these two data sets is that data set I considers the week along with soil type, but data set II calculates the day wise pesticides used. With this data we can cover and find all over pesticide usage. Other data we considered is pesticide data over years and dates. This dataset III that tells crop is alive or damaged by using pesticides levels with water sheds. Water shed is a water capacity holding storage that describes the water quantity level in each crop.

By using all these we finally collected all factors that can affect the crop life. So majorly what we had observed is crop is affected due to pesticide dosage levels and this severity levels will depend upon pesticides dosage over days and over week with respect to states.

Random Forest Classifier Algorithm and Decision Tree Classifier Algorithm given correct prediction and results. So, by implementing these algorithms we can save our crops which is a need for every human. By Tableau we analyzed the pollutant effect rate. Cat Boosting Classifier is used with 1000 iterations along with the depth of 7.

As we observed the pattern that crop life will majorly affect due to dosage and number of weeks used, it can be easily observed and can know by using Multi-Layer Perceptron Classifier. The relation could be a Linear relation such that if pesticides dosages are greater than 67 then crop is going to be dead. This is a main case we had observed, so we can say that it is a positive correlation between these two attributes. Decision Tree Classifier is a Supervised Learning algorithm that constructs a tree based on Gini Index by calculating the entropy it

splits the node. Here it is splitting the node (Main /Root) Pesticides Category. Similarly, Random Forest Classifier internal working include Decision Tree Classifier. But we had made into 100 iterations. And provide depth as 8. Cat Boost is an Categorical Boosting Classifier that gives boost to all the attributes to increase the accuracy. We are providing the depth as 7 and learning rate as 0. 1 and iterating over 1500 times.

Table 9: ML Algorithm Analysis on Datasets

Algorithm	**Data I**	**Data II**	**Data III**
Random Forest Classifier	97. 99%	85. 08%	89. 67%
Bagging Classifier	96. 33%	85%	87. 74%
Ada Boost Classifier	84. 08%	84. 53%	84. 33%
LGBM Classifier	85. 03%	86. 05%	86. 75%
Naïve Bayesian Classifier	80. 63%	84. 16%	74. 16%
Decision Tree Classifier	97. 99%	84. 58%	84. 98%
Cat Boost Classifier	85. 98%	84. 73%	85. 13%
Perceptron	80. 38%	85. 03%	87. 56%
Multi Layer Perceptron	83. 60%	83. 99%	79. 67%

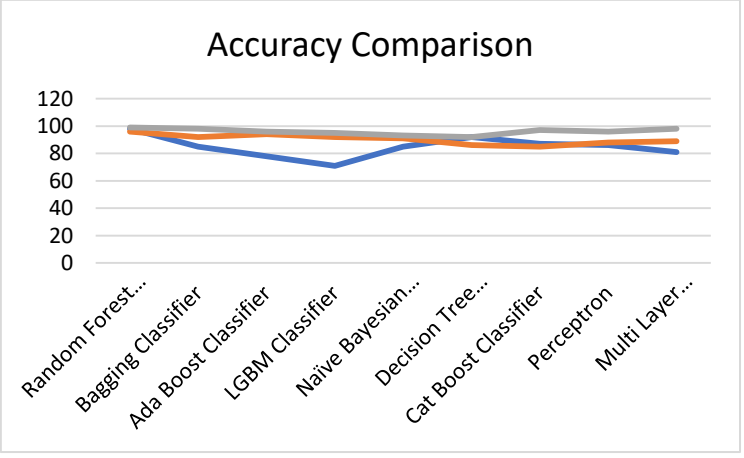

Figure: 1 Accuracy Comparison

Here we are going to convert Light GBM Classifier Model into a pickle file to make an web application. Flask is an web application framework supported by python. We are going to deploy our Paper into Flask, Such that Training Data is given as input with HTML Textboxes. Finally, flask application is dumped in an EC2 Amazon Web Server. This application is achieved by using Android Application which includes button widgets to get the predicted value information. It also includes other information about pesticide dose calculation for calculating pest dosages used in week days. Compost pit is calculated based on No of rows and columns in a field. Images are given in the app for crop information. Pests and crops clarifications are clarified by scientist which is supported by our app. Fig – 3. 10. 2. 1 shows the Main view of our app. Crop Details include crop images along with the information. View pest details tells for which cop what pesticide is used on daily basis. Our App has another feature that trace the location and tells which crop gets high yield and productivity based on climate factors in crop – season details. Seasons are given along with districts for better accuracy. Our Flask web server application IP address is mentioned in "click here to predict crop" button which implements an Onclick Property such that we can move on to predicting the crop life. This app is created by using Java language in Android Studio. User Interface Design is created using Extensible markup Language (XML). Fig 4. 2 shows the flask web application to predict the crop life.

5. Conclusions and Future Scope of Work

Our Main enhancement is to decrease the pesticides and insecticides so that crop can alive for different seasons and give necessary suggestions for crops. From the above work, I conclude that Random Forest Classifier had given the best accuracy for predicting the crop life, Mobile application can be build to help farmers to know the pesticides to relevant crops. It also calculates the moisture content for each crop. By using this app, Crop irrigation can be achieved for farms to get higher yield. We are also providing the information on the disadvantages of using large amount of pesticides and introduced the usage of compost pit which is not poisonous. The feature extraction had given the best scope for calculating the pesticide dosage. If any problems or clarifications on pesticides our app provides scientist contact details such that farmers can contact directly by using this app.

References

[1] Pavan Patil, Virendra Panpatil, Prof. Shrikant, Crop Prediction System using Machine Learning Algorithms Issue 02 Feb 2020, IRJET.

[2] Cao Xiangliang, Fu Han-Chi, Anthony Lam. Agricultural product monitoring system supported by cloud computing. 2018, SpringerScience+BusinessMedia.

[3] Marzieh Mokarram, Mohammad R, Khosravi. A cloud computing framework for analysis of agricultural big data based on Dempster – Shafer theory. 2020, Springer Science+Business Media.

[4] SanazRasti, Chris J. Bleakle, Gue´nole´ C. M, Silvestre N. M. Holden David Langton, Gregory M. P. Crop growth stage estimation prior to canopy closure using deep learning algorithms. 2020, Springer.

[5] ShinwooHyun, Seung Mo Yang, Junhwan Kin c. Development of a mobile computing framework to aid decision-making on organic fertilizer management using a crop growth model. 2020, IRJET.

[6] A. S. Gardner. I. M. D. Maclean, K. J. Gaston. Forecasting future crop suitability with microclimate data. 2019, Springer.

[7] Ekaansh Khosla, Ramesh Dharavath, Rashmi Priya. Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression. 2019, Springer.

[8] Shital H, Nirav Bhatt. Wheat crop yield prediction using new activation function in a neural network. 2020, Springer.