International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT228339

Agricultural Production Optimization

A study to maximize the agricultural production and minimize the use of fertilizers

Kavya Gupta[1], Nidhi Senger[2], Vasudha Bahl[2], Amita Goel[2]

[1]Student, Department of Information Technology, Maharaja Agrasen Instute of Technology, GGSIP University, Delhi, India

[2]Faculty, Department of Information Technology, Maharaja Agrasen Instute of Technology, GGSIP University, Delhi, India

Article Info
Volume 8, Issue 3
Page Number : 135-141

Publication Issue :
May-June-2022

Article History
Accepted: 10 May 2022
Published: 22 May 2022

ABSTRACT

A large number of people have attempted to create mathematical models in order to develop and optimize agricultural production; however almost none of them address the factors such as use of fertilizers (N, K, P) and PH of the soil. More use of fertilizers can deteriorate the fertility of soil and can lead to the low production for years to come. In this study a predictive model is formulated that aims to predict the best crop to be grown in available weather conditions and nutrients present in the soil thereby reducing the use of chemicals and taking a step forward towards organic and chemical free agriculture. This model also focuses on the point that there is no need to grow the same crop in same field every time.

Keywords: Optimize Agricultural Production, Fertilizers

I. INTRODUCTION

Agriculture plays a vital role in the global economy. India is home to more than 130 crores people and to cater to their food demand is a major challenge. According to FAO estimates in 'The State of Food Security and Nutrition in the World, 2020 report, 189.2 million people are undernourished in India. By this measure 14% of the population is undernourished in India. With the ever increasing human population, understanding worldwide crop yield is central to addressing food security challenges and reducing the use of fertilizers. The demand for food crops have been increasing but is very difficult to increase the area of land under agriculture with so much pressure on land.[3] So there should be a methodology that increases the crop production on the same amount of land that is available with us. Crop yield prediction is an important agricultural problem. The Agricultural yield primarily depends on weather conditions (rain, temperature), pesticides and water availability. Accurate information about history of crop yield is important to get the maximum output of crop yield on the same dimensions of available land and to reduce the use of fertilizers so we can move towards organic farming. This problem not only persist in India but Worldwide as estimates show that, in 2020, over 237 crore people were grappling with food insecurity globally, an increase of about 32 crores from 2019. With such a negative report there should be some step taken to overcome the above problems. When you don't rotate crops, as farmers thousands of

years ago discovered, a slew of issues arise. Over the course of several years, all of these issues can result in lower yields.

To begin with, the soil may grow "weary" and less fruitful. This is due to the same type of crop being planted in the same place over and over, depleting the land of the same nutrients required for the plant's growth. Second, when pests learn to create a home near a field that constantly produces

the same type of crop, they might reach levels that are difficult to control. Finally, if the same sort of crop is grown season after season, the soil may be more sensitive to erosion effects.

Crop rotation is beneficial. On the same grounds, data was studied that There are three major nutrition requirements for any crop to grow that is nitrogen, phosphorus and potassium also known as NPK. we extended the methodology of [1] and [2] so that a predictive model can be created that majorly work on 8 factors namely Nitrogen ratio in soil, phosphorous ratio in soil, potassium ratio in soil, pH of the soil, rainfall, temperature and humidity. With the study of all this factors a predictive model is created that can predict which crop can be grown best with available weather and nutritional value of soil. This model is based on the pre- existing factors that urges farmers to reduce their dependency on external factors such as using of chemicals and rainfall and being more dependent on what they already have. After the testing of soil in lab, NPK ratio and pH can be found out easily and this model can be applied. This model will suggest the crop based on its predictive model that should be grown on the field so there is less use of fertilizers and other chemicals

II. METHODOLOGY

This paper develops a predictive model that effectively addresses the above problem. This model not only considers the weather and climatic condition but also the NPK ratio and pH of the soil

A. DATA SET

Data set has been created by going through the data for agriculture through various website

- https://www.fao.org/3/a0257e/A0257E05.htm
- https://agricoop.nic.in/sites/default/files/Horticulture%20Statistics%20at%20a%20Glance-2018.pdf

Dataset contains 3200 rows containing weather requirement and NPK requirement for different crops and different fields with 8 columns namely N- Ratio of nitrogen content in the soil, P- Ratio of phosphorus content in the soil, K- Ratio of potassium content in the soil, Temperature- Temperature required for the crop to grow in degree Celsius, Humidity- Relative humidity in percentage, pH- pH value of the soil, Rainfall- Rainfall in mm and Label- Name of different crops This dataset consists the data for various fields from all over India. in all total 32 crops were taken and for each crop 100 observation were taken as follows Kidney beans, cotton, banana, orange, coconut, moth beans, chickpea, mango, mung bean, lentil, pigeon peas, apple, rice, muskmelon, pomegranate, coffee, maize, black gram, jute, papaya,

grapes, watermelon, potato, onion, black bean, soya bean, Ragi, wheat. All the 32 crops data was accumulated in the form of excel table and was imported in python project so that predictive model can be created. The same data was divided later on into testing and training data. 255 rows were the part of testing data and 75% rows were the part of training data. There was equal distribution of crops

	N	P	K	temperature	humidity	ph	rainfall	label
0	90	42	43	20.879744	82.002744	6.502985	202.935536	rice
1	85	58	41	21.770462	80.319644	7.038096	226.655537	rice
2	60	55	44	23.004459	82.320763	7.840207	263.964248	rice
3	74	35	40	26.491096	80.158363	6.980401	242.864034	rice
4	78	42	42	20.130175	81.604873	7.628473	262.717340	rice

Given above is the snippet of dataset that shows the first 5 rows .it conatins the data for rice that is grown in different parts of india and gives the information

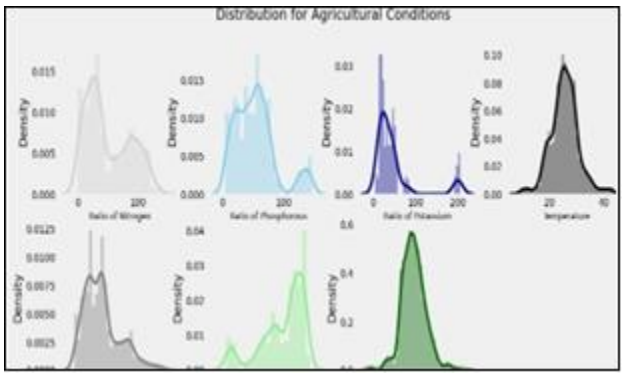
regarding N, P, K, temprature, humidity, ph and rainfall.

B. DESCRIPTIVE STATISTICS

The average ratio of nitrogen, phosphorous and also the average temperature, humidity, pH value and rainfall for each crop was checked. After the summary statistics for individual crop was found, an average function was used and minimum, maximum and average for each factor for each crop was found. It was inferred that rice requires good rainfall and banana requires good proportion of potassium and so on for all other crops. After this average requirement of each crop was compared with average conditions as well using interactive functions. Some crops require less nitrogen such as coconut, apple and some require more nitrogen such as cotton, muskmelon etc and such comparisons were taken for each factor.After this crops which require more than or less than the average given condition to grow were found.

C. ANALYZING AGRICULTURE CONDITION

Distribution of different agricultural conditions were found using distplot function of seaborn library of python. Below is the screenshot of the graph obtained



After studying th above graphs, potassium and humidity factor were found to be different. Potassium graph was right

skewed and humidity was left skewed and all other conditions are normally distributed. Next step was to look for particular crops were found that requires high and low amount of NPK and other factors and after that Crops were classify on the basis of season summers, winters and rainy on the basis of temperature and humidity filters. pH graph was neither right skewed nor left skewed. This shows that most of the crops require an average amount 6-7 pH range of the soil for the best crop to be grown. Still right amount of pH becomes an important deciding factor for the crop to be grown in perfect conditions.

D. CLUSTERING

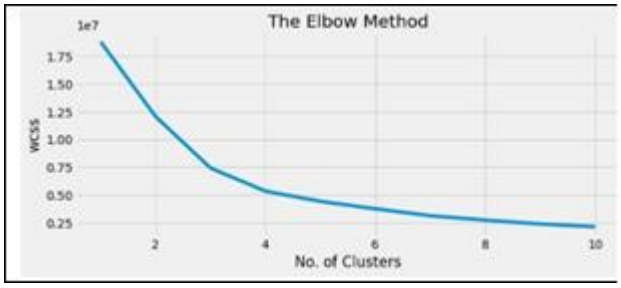
In clustering target number k is defined, which refers to the number of centroids needed in the dataset. A centroid is the imaginary or real location representing the centre of the cluster. Every data point is allocated to each of the clusters through reducing the in-cluster sum of squares. In other words, the K-means algorithm identifies k number of centroids, and then allocates every data point to the nearest cluster, while keeping the centroids as small as possible.

It is process of grouping data points with similar traits and assigning them to clusters. K means clustering is used to determine the optimal number of cluster to be made elbow method was used. Unsupervised Learning algorithm K- Means Clustering divides the unlabeled dataset into various clusters. K specifies the number of pre-defined clusters that must be produced during the process; for example, if K=2, two clusters will be created, and if K=3, three clusters will be created, and so on. It allows us to cluster data into different groups and provides a simple technique to determine the categories of groups in an unlabeled dataset without any training.

It's a centroid-based approach, which means that each cluster has its own centroid. The main goal of this technique is to reduce the sum of distances between data points and the clusters that they belong to.

The algorithm starts with an unlabeled dataset and divides it in half.Clusters of crop were made based on

the factors from data set so that they can be grouped and a predictive model can be created.



The average is between 3 and 4 so 4 clusters were considered. KMeans algorithm was implemented and 4 clusters were as follows.

But some of the crops were in more than one clustering so hard clustering was done instead of soft clustering so that a crop can be contained in one cluster only.

Results for Hard Clustering

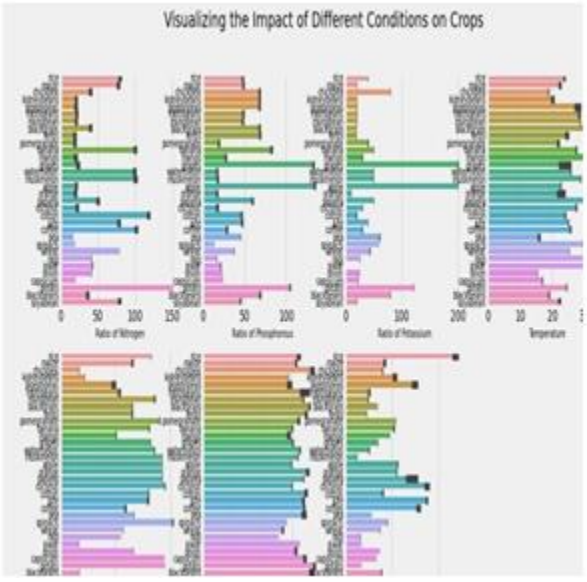
Crops in Cluster 1: ['banana', 'wheat', 'muskmelon', 'soyabean', 'maize', 'potato', 'watermelon', 'cotton']

Crops in Cluster 2: ['pigeonpeas', 'rice', 'coconut', 'papaya', 'coffee', 'jute']

Crops in Cluster 3: ['mothbeans', 'spinach', 'ragi', 'capsicum', 'blackgram', 'mungbean', 'lentil', 'onion', 'orange' 'pomegranate', 'blackbeans', 'pea', 'mango', 'kidneybeans', 'chickpea']

Crops in Cluster 4: ['grapes', 'apple']

E. SOME PATTERNS



The above graph was created to find out some patterns like cotton require high quantity of nitrogen and grapes and apple requires high amount of phosphorus and other factors for other crops

F. PREDICTIVE MODEL

Dataset was split into independent and dependent dataset. Dependent variable was label column(crop) Later it was split into training and testing data. Finally, model was created using logistic regression. Logistic regression was imported from sklearn.linear model and the object was created and was training data set was fil into the model

```
# create a Predictive Model

from sklearn.linear_model import LogisticRegression
model = LogisticRegression()
model.fit(x_train, y_train)
y_pred = model.predict(x_test)
```

III. RESULT

After creating the model real time data was used in which a person has to enter N RATIO,P RATIO, K RATIO, RAINFALL, HUMIDITY, PH , TEMPRATURE so the model can predict which crop can be best grown in the available conditions Below case predicts that rice is the most suitable crop to be grown with given conditions where N=90 , P=40 , K=20, temperature=15 , humidity=80, ph=7, rainfall=200cm.

```
prediction =model.predict((np.array([[90,
                                      40,
                                      20,
                                      15,
                                      80,
                                      7,
                                      200]])))
print("The Suggested Crop :", prediction)
```

Test dataset was predicted with the help of predict function

```
prediction =model.predict((np.array([[19,
                                      45,
                                      64,
                                      13,
                                      68,
                                      5,
                                      55]])))
print("The Suggested Crop :", prediction)
```

```
The Suggested Crop : ['pea']
```

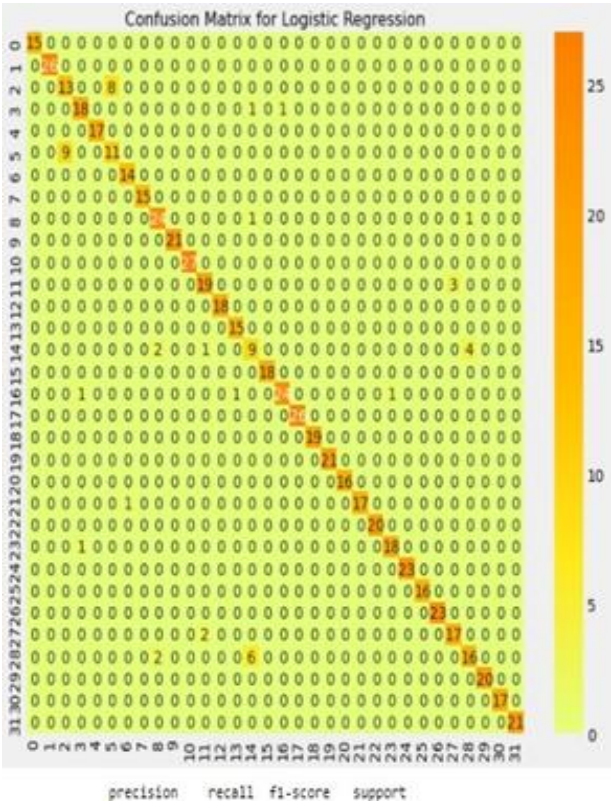
Fig2: result 2

```
prediction = model.predict((np.array([[90,
                                       40,
                                       40,
                                       20,
                                       80,
                                       7,
                                       200  ]])))
print("The Suggested Crop for Given Climatic Condition and the amount of NPK is :", prediction)
```

```
The Suggested Crop for Given Climatic Condition and the amount of NPK is : ['maize']
```

Fig3: result 3

Heat map and confusion matrix was created and Good results for precision recall and f1-score was found and the accuracy was about 93% for the training data set which is easily visible in the below snippet of heat map.



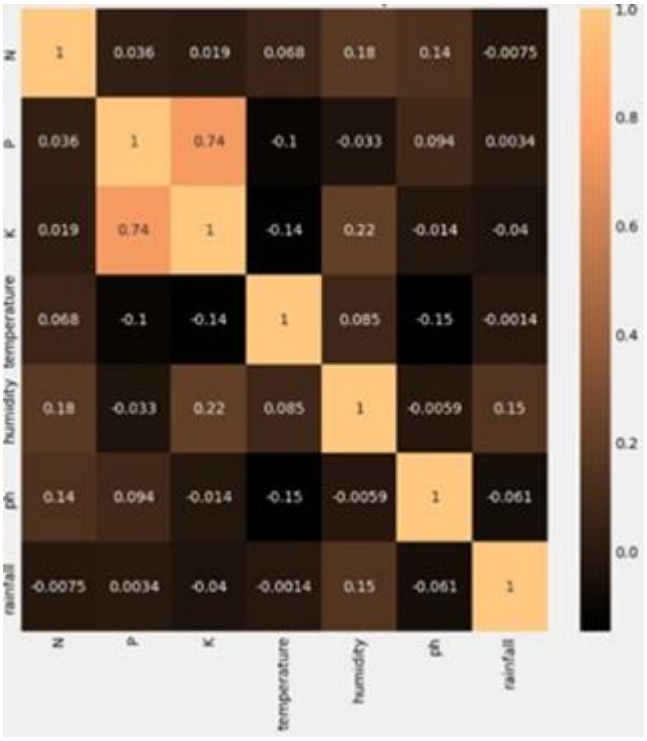
rice	0.85	0.89	0.87	19
soyabean	0.76	0.67	0.71	24
spinach	1.00	1.00	1.00	20
watermelon	1.00	1.00	1.00	17
wheat	1.00	1.00	1.00	21
accuracy			0.93	640
macro avg	0.93	0.93	0.93	640
weighted avg	0.93	0.93	0.93	640

Particular crops that require less NPK and the crops that require more of rainfall and temperature were fo found

```
In [334]: data[(data['temperature'] > 40) & (data['humidity'] > 20) & (data['rainfall'] > 200)]['label'].unique()
Out[334]: array(['papaya'], dtype=object)

In [336]: data[(data['N'] < 10) & (data['P'] < 10) & (data['K'] < 10)]['label'].unique()
Out[336]: array(['orange'], dtype=object)
```

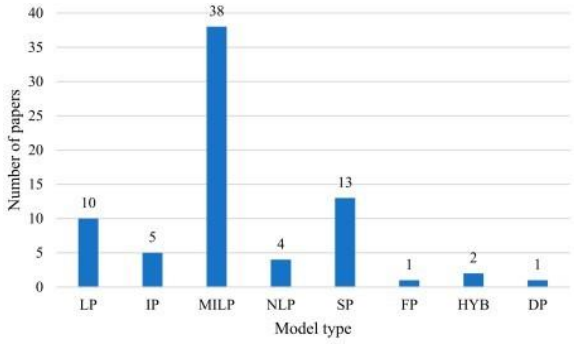
After this heatmap function was used to find the correlation



Predictive modelling was compared to all the other available models

LP: Linear programming, IP: Integer programming, MILP: Mixed-integer linear programming, NLP: Nonlinear Programming, SP: Stochastic programming, FP: Fuzzy programming) HYB stands for hybrid, and DP stands for dynamic programming. MILP was the most popular modelling method (e.g., [47,89]). Notably, linear programming and integer programming models presented a pure integer non-

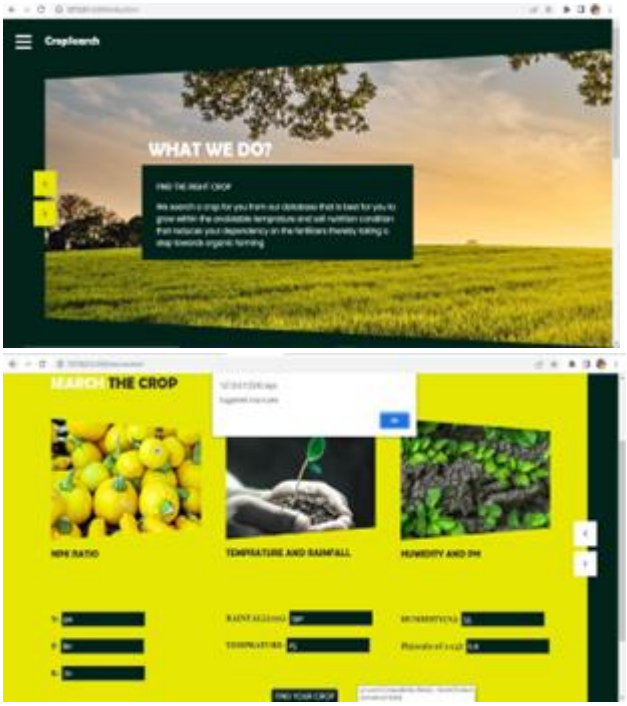
linear programming model to create a green harvest scheduling A multi-objective mixed-integer non-linear programming paradigm was proposed by AFSC. Bi-level optimization was explored as a way to model a non- linear programming model and create a durable optimization model. To deal with uncertainties in AFSCs, a number of articles used various types of stochastic programming in their models. Two-stage stochastic programming, resilient optimization, and scenario-based stochastic programming, for example, have all been investigated. There are only two hybrid models.] created a multi-criteria method that combines the Analytic Hierarchy Process and Ordered Weighted Averages. A novel hybrid stochastic fuzzy-robust programming approach was given, as well as four other hybrid approaches, by averaging to evaluate stakeholders and using the answer as an input parameter in the mathematics.



Below is the data set link on which predictive model was created that consist of 3200 rows occupying that data of 32 different crops

https://drive.google.com/file/d/1oyyH-LpJiSy9_tDmEkCy6Z0itEU3-Awt/view?usp=sharing

This predictive model was parsed to the JSON format and then UI-UX was created. Below are the screenshots attached for the website



IV. CONCLUSION

This help the farmer to grow the right crop with given condition and as that NPK requirement will be fulfilled this will considerably reduce the use of chemical fertilizers and improve the quality of food and perfect climatic condition will reduce the chances of crop failure and hence improving the food security for the country. Basic idea for this model is that there is no need to grow the same crop in the same field every time. You just need to check the soil conditions and this model will help to find the right crop for your field and reduce the dependency on the fertilizers.

II. REFERENCES

[1]. Basil Manos, Parthena Chatzinikolaou, Fedra Kiomourtzi, Sustainable Optimization of Agricultural Production, APCBEE Procedia, Volume 5, 2013, Pages 410-415,ISSN 2212-6708,.published

[2]. Agricultural Production Structure Optimization: A Case Study of Major GrainProducing Areas, ChinaLU Sha- sha1, 2, LIU Yan-sui1, LONG Hua-lou1 and GUAN Xing-liang1, published

[3]. S. Saranya and T. Amudha, "Crop planning optimization research — A detailed investigation," 2016 IEEE International Conference on Advances in Computer Applications (ICACA), 2016, pp. 202-208, doi: 10.1109/ICACA.2016.788795

[4]. F M Zenis , S Supian , E LesmanaIOP Conf. Series: Materials Science and Engineering 332 (2018) 012021 doi:10.1088/1757-899X/332/1/012021

[5]. J. Brenes and G. Raventos, "Research on Optimization Techniques Applied to Precision Agricultural Processes Supported by Technology in Latin America: Opportunities and a Route to Follow," in 2020 XLVI Latin American Computing Conference (CLEI), Loja, Ecuador, 2020 pp. 462-470. doi: 10.1109/CLEI52000.2020.00060 https://doi.ieeecomputersociety.org/10.1109/CLEI5200 0.2020.00060

[6]. Adams, Darius M.; Alig, Ralph J.; Callaway, J.M.; McCarl, Bruce A.; Winnett, Steven M. 1996. The forest and agricultural sector optimization model (FASOM): model structure and policy applications. Res. Pap. PNW-RP-495. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 60 p

[7]. Uguru, H., and Akpenyi-Aboh, O. N. (2021).

[8]. Optimization of agricultural machines through the pre- harvest treatment of sweet paper (cv. Goliath) fruits. Direct Research Journal of Agriculture and Food. Science.Vol. 9, Pp. 167-173

Cite this article as :

Kavya Gupta, Nidhi Senger, Vasudha Bahl, Amita Goel, "Agricultural Production Optimization", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 3, pp. 135-141, May-June 2022. Available at doi : https://doi.org/10.32628/CSEIT228339

Journal URL : https://ijsrcseit.com/CSEIT228339