# ARMANDO MALUANA, MSISM

GILBERT, AZ

(480) 745-4762 • armmaluana@gmail.com • LinkedIn

# **Agricultural Production Optimization**

**Problem Statement:** Construct a Predictive Machine Learning Model to enhance crop selection based on available climate and soil conditions.

**Goal:** Enhance Precision Farming practices by optimizing agricultural processes.

This Project Focuses on Advancing Precision Farming:

- 1- Maximizing Productivity
- 2- Deepening Understanding of Climate and Soil Requirements for Crops
- 3- Assisting Users in Dealing with Unpredictable Weather Patterns

**Tools:** The Machine Learning model was developed using Jupyter Notebook in Python. The toolkit includes essential libraries such as Pandas, NumPy, Matplotlib, Seaborn, and Interact.

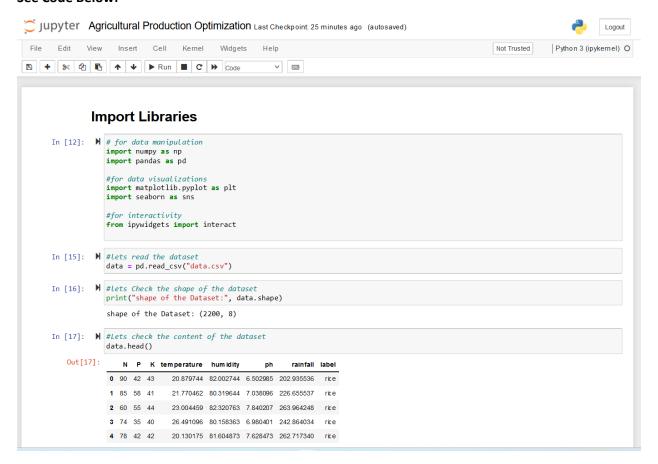
The dataset encompasses a diverse range of crops, each with specific climate and soil preferences. I crafted feature vectors and divided the data into 80% for training and 20% for testing purposes. With the model now well-trained, users can input values for Nitrogen (N), Phosphorus (P), and Potassium (K), along with Temperature, Humidity, pH, and Rainfall.

Based on the input data, the model furnishes recommendations for the most suitable crops, accounting for the prevailing soil and climatic conditions.

#### Final Outcome:

•

## See Code Below:



```
In [18]: N #lets check if there is any missing value in the dataset
#if there missing values you can use the fillin function. you can use:
##lean-numerical values,
#Median - If the calues are the dataset contains large numbers of utliers.
##lode-If the values are categorrical
data.isnull().sum()
       Out[18]: N
                        temperature
                        humidity
                       ph
rainfall
                       label
dtype: int64
In [21]: M #Lets check the Crops present in the Dataset
data["label"].value_counts()
       Out[21]: rice
                                                 100
100
                       maize
                       jute
cotton
                                                  100
                                                 100
100
                        coconut
                       papaya
orange
apple
muskmelon
                                                  100
                                                  100
                                                  100
                        watermelon
                                                  100
                       grapes
mango
                                                  100
                                                  100
                                                 100
100
                        banana
                        pomegranate
lentil
                                                  100
                                                 100
100
                        blackgram
                        mungbean
                        mothbeans
                                                  100
                        pigeonpeas
                                                  100
```

```
print("\033[1m" + "Statistics for Rainfall" + "\033[0m")
print("Minimum Rainfall required :", X["rainfall"].min())
print("Minimum Rainfall required :", X["rainfall"].mean())
print("Maximum Rainfall required :", X["rainfall"].max())
        crops rice
 rice
Statistics maize
Minimum Nit Jute
Minimum Nit cotton
Maximum Nit coconut
     ----- papaya
Statistics orange
Minimum Phc apple
Minimum Pho
muskmelon
Maximum Pho watermelon
Statistics grapes
Minimum Pot mango
Minimum Pot banana
Maximum Pot pomegranate
             -- lentil
Statistics blackgram
                                                          142
Minimum Tem mungbean
Minimum Tem mothbeans
                                                          32210500005
Maximum Tem pigeonpeas
                                                          5077
Statistics for Humidity
Minimum Humidity required: 80.12267476
Minimum Humidity required: 82.2728215389
Maximum Humidity required : 84.96907151
Statistics for PH
Minimum PH required : 5.005306977
Minimum PH required : 6.425470922139999
Maximum PH required : 7.868474653
```

```
In [13]: | #lets make this function more interactive @interact

def compare (Conditions = ["N","P","K","temperature","humidity","ph","rainfall"]):
    print("\033[1m" +"Crops which require greater than average", Conditions,"\n" + "\033[0m")
    print(data[data[Conditions] > data[Conditions].mean()]["label"].unique())
    print("\033[1m" +"Crops which require less than average", Conditions,"\n" + "\033[0m")
    print(data[data[Conditions] <= data[Conditions].mean()]["label"].unique())

Conditions lemperature

['rice' 'maize' 'pigeonpeas' 'mothbeans' 'mungbean' 'blackgram' 'lentil'
    'banana' 'mango' 'grapes' 'watermelon' 'muskmelon' 'orange' 'papaya'
    'coconut' 'cotton' 'jute' 'coffee']

Crops which require less than average temperature

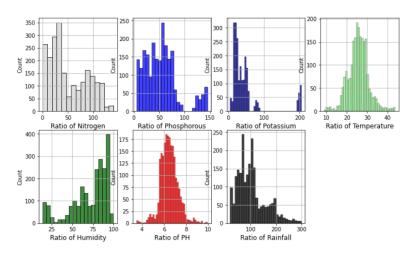
['rice' 'maize' 'chickpea' 'kidneybeans' 'pigeonpeas' 'mothbeans'
    'blackgram' 'lentil' 'pomegranate' 'banana' 'grapes' 'watermelon' 'apple'
    'orange' 'papaya' 'coconut' 'cotton' 'jute' 'coffee']
```

## **Distribution**

```
In [117]: # # Lets check the Distribution of Agricultural Conditions of dataset
    plt.rcParams["figure.figsize"] = (12,7)

    plt.subplot(2, 4, 1)
    sns.histplot(data["N"], color = "lightgrey")
    plt.subplot(2, 4, 2)
    sns.histplot(data["P"], color = "blue")
    plt.subplot(2, 4, 2)
    sns.histplot(data["P"], color = "blue")
    plt.subplot(2, 4, 3)
    sns.histplot(data["K"], color = "darkblue")
    plt.subplot(2, 4, 3)
    sns.histplot(data["K"], color = "darkblue")
    plt.subplot(2, 4, 4)
    sns.histplot(data["Kemperature"], color = "lightgreen")
    plt.subplot(2, 4, 4)
    sns.histplot(data["kemperature"], color = "lightgreen")
    plt.subplot(2, 4, 5)
    sns.histplot(data["humidity"], color = "darkgreen")
    plt.subplot(2, 4, 5)
    sns.histplot(data["humidity"], color = "darkgreen")
    plt.subplot(2, 4, 6)
    sns.histplot(data["humidity", fontsize = 12)
    plt.grid()

plt.subplot(2, 4, 6)
    sns.histplot(data["humidity", fontsize = 12)
    plt.grid()
```



```
In [15]: M

# Lets find out some Interesting Facts

print("Some Interesting Patterns")

print("Crops which require very High Ratio of Nitrogen Content in Soil:", data[data["N"]:120]["label"].unique())

print("Crops which require very High Ratio of Potassium Content in Soil:", data[data[""]:200]["label"].unique())

print("Crops which require very High Ratio of Potassium Content in Soil:", data[data[""]:200]["label"].unique())

print("Crops which require very High Ratio fill:", data[data["rainfall"]:200]["label"].unique())

print("Crops which require very Low Temperature:", data[data["rainfall"]:200]["label"].unique())

print("Crops which require very High Temperature:", data[data["temperature"]:40]["label"].unique())

print("Crops which require very High PH:", data[data["ph"]:9]["label"].unique())

print("Crops which require very High Ratio of Nitrogen Content in Soil: ['cotton']

Crops which require very High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very Low High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very Low High Ratio of Potassium Content in Soil: ['grapes' 'apple']

Crops which require very Lo
```

```
In [16]:  # Lets understand which crops can only be Grown in Summer, Whiter and Rainy Season

print ("\033[1m" +"Summer Crops"+ "\033[0m")

print(data[(data["temperature"]>30)&(data["humidity"]>50)]["label"].unique(), "\n")

print ("\033[1m" +"Winter Crops"+ "\033[0m")

print(data[(data["temperature"]<20)&(data["humidity"]>30)]["label"].unique(), "\n")

print ("\033[1m" +"Rainy Crops"+ "\033[0m")

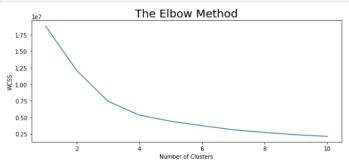
print(data[(data["rainfall"]>200)&(data["humidity"]>30)]["label"].unique())

Summer Crops
['pigeonpeas' 'mothbeans' 'blackgram' 'mango' 'grapes' 'orange' 'papaya']

Winter Crops
['maize' 'pigeonpeas' 'lentil' 'pomegranate' 'grapes' 'orange']

Rainy Crops
['rice' 'papaya' 'coconut']
```

```
In [18]: ▶ #Lets Cluster Crops
              #Lets import the warning libraries so that we can avoid warning
              from sklearn.cluster import KMeans
              import warnings
warnings.filterwarnings("ignore")
              #lets select Colums from the Dataset x = data.loc[:,["N","P","K","temperature","humidity","ph","rainfall"]].values
              #Let's check the shape of x
              print(x.shape)
               #Lets convert this data into a dataframe
              x_data = pd.DataFrame(x)
x_data.head()
              (2200, 7)
    Out[18]:
                    0
                        1
                              2
                                        3
                                                  4
                                                           5
               0 90.0 42.0 43.0 20.879744 82.002744 6.502985 202.935536
               1 85.0 58.0 41.0 21.770462 80.319644 7.038096 226.655537
               2 60.0 55.0 44.0 23.004459 82.320763 7.840207 263.964248
               3 74.0 35.0 40.0 26.491096 80.158363 6.980401 242.864034
               4 78.0 42.0 42.0 20.130175 81.604873 7.628473 262.717340
```



```
In [20]: M #Next, we are going to implement the K Means algorithm to perform clustering analysis km = KMeans(n_clusters = 4, init = "k-means++", max_iter = 300, n_init = 10, random_state = 0)
                y_means = km.fit_predict(x)
                #Lets find out the results
a = data["label"]
                y_means = pd.DataFrame(y_means)
                z = pd.concat([y_means, a], axis = 1)
z = z.rename(columns ={0:"cluster"})
                 #Lets check the Clusters of each Crops
                print("Lets check the Results After Applying the K Means Clustering Analysis \n") print("Crops in First Cluster:", z[z["cluster"] == \theta]["label"].unique())
                 print("--
                print("Crops in Second Cluster:", z[z["cluster"] == 1]["label"].unique())
                 print("Crops in Third Cluster:", z[z["cluster"] == 2]["label"].unique())
                print('
                print("Crops in Fourth Cluster:", z[z["cluster"] == 3]["label"].unique())
                 Lets check the Results After Applying the K Means Clustering Analysis
                Crops in First Cluster: ['maize' 'chickpea' 'kidneybeans' 'pigeonpeas' 'mothbeans' 'mungbean' 'blackgram' 'lentil' 'pomegranate' 'mango' 'orange' 'papaya' 'coconut']
                 Crops in Second Cluster: ['maize' 'banana' 'watermelon' 'muskmelon' 'papaya' 'cotton' 'coffee']
                 Crops in Third Cluster: ['grapes' 'apple']
                 Crops in Fourth Cluster: ['rice' 'pigeonpeas' 'papaya' 'coconut' 'jute' 'coffee']
```

```
In [21]: 
| Hets build a predictive Module to help farmers deide the type of Crops given climate and soil conditions.

##Lets split the Dataset from Predictive Modeling

y = data["label"]
 x = data.drop(["label"], axis = 1)

print("Shape of x:", x.shape)

Shape of x: (2200, 7)

Shape of y: (2200,)

In [22]: 
| #Now we are creating Training and Testing Sets for Validation of Results
from sklearn.model_selection import train_test_split

x_train, x_test, y_train,y_test = train_test_split(x,y, test_size = 0.2, random_state =0)
print("The Shape of x train:", x_train.shape)
print("The Shape of x train:", x_train.shape)
print("The Shape of y train:", y_train.shape)
print("The Shape of y test:", y_test.shape)

The Shape of x train: (1760, 7)
The Shape of y test: (440, 7)
The Shape of
```

```
In [175]: N
#Lets evaluate the Model Performance, we see that the Recall in 97% this sugests we are accurate
from sklearn.metrics import classification_report

#Lets print the classification Report as well
#this table gives us the accuracy of the model, greater value of the Precision and Recall = better the Model
cr = classification_report(y_test,y_pred)
print(cr)
```

	precision	recall	f1-score	support
apple	1.00	1.00	1.00	18
banana	1.00	1.00	1.00	18
blackgram	0.86	0.82	0.84	22
chickpea	1.00	1.00	1.00	23
coconut	1.00	1.00	1.00	15
coffee	1.00	1.00	1.00	17
cotton	0.89	1.00	0.94	16
grapes	1.00	1.00	1.00	18
jute	0.84	1.00	0.91	21
kidneybeans	1.00	1.00	1.00	20
lentil	0.94	0.94	0.94	17
maize	0.94	0.89	0.91	18
mango	1.00	1.00	1.00	21
mothbeans	0.88	0.92	0.90	25
mungbean	1.00	1.00	1.00	17
muskmelon	1.00	1.00	1.00	23
orange	1.00	1.00	1.00	23
papaya	1.00	0.95	0.98	21
pigeonpeas	1.00	1.00	1.00	22
pomegranate	1.00	1.00	1.00	23
rice	1.00	0.84	0.91	25
watermelon	1.00	1.00	1.00	17
accuracy			0.97	440
macro avg	0.97	0.97	0.97	440
weighted avg	0.97	0.97	0.97	440

```
In [24]: 🔰 #Now we are done with testing, now we are going to do predictions. First we check the head of the dataset
             data.head()
    Out[24]:
                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
             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
43,
                                                 30,
100,
            8,
100]])))
print("The Sugested Crop for Given Climatic Condition is :", prediction)
             The Sugested Crop for Given Climatic Condition is : ['jute']
```