

# **ESTIMATION OF CROP YIELD USING DATA ANALYTICS**

## **1. INTRODUCTION**

### **1.1 Project Overview**

Crop production in India is one of the most important sources of income and India is one of the top countries to produce crops. As per this project we will be analyzing some important visualization, creating a dashboard and by going through these we will get most of the insights of Crop production in India.

### **1.2 Purpose**

Crop yield prediction is an essential task for the decision-makers at national and regional levels for rapid decision-making. An accurate crop yield prediction model can **help farmers to decide on what to grow and when to grow**. There are different approaches to crop yield prediction.

## **2. LITERATURE SURVEY**

### **2.1 Existing problem**

#### **ABSTRACT:**

Agriculture is important for human survival because it serves the basic need. A well-known fact that the majority of population ( $\geq 55\%$ ) in India is into agriculture. Due to variations in climatic conditions, there exist bottlenecks for increasing the crop production in India. It has become challenging task to achieve desired targets in Agri based crop yield. Factors like climate, geographical conditions, economic and political conditions are to be considered which have direct impact on the production, productivity of the crops. Crop yield prediction is one of the important factors in agriculture practices. Farmers need information regarding crop yield before sowing seeds in their fields to achieve enhanced crop yield. The use of technology in agriculture has increased in recent year and data analytics is one such trend that has penetrated into the agriculture field being used for management of crop yield and monitoring crop health. The recent trends in the domain of agriculture have made the people to understand the significance of Big data. The main challenge using big data in agriculture is identification of impact and effectiveness of big data analytics.

### **2.2 References**

**TITLE :** Agriculture Data Analytics in Crop Yield Estimation: A Critical Review.

**AUTHOR :** B.M.Sagar, Cauvery N K

**TITLE :** Machine learning approach for forecasting crop yield based on climatic parameters

**AUTHOR :** S.Veenadhari, Dr. Bharat Misra&Dr. CD Singh

**TITLE :** A Survey on Crop Yield Prediction based on Agricultural Data

**AUTHOR :** Dhivya B H, Manjula R, Siva Bharathi S, Madhumathi R

## TITLE : ANALYSIS OF CROP YIELD PREDICTION USING DATA MINING TECHNIQUES

AUTHOR :D Ramesh ,B Vishnu Vardhan

### 2.3 Problem Statement Definition

Create a problem statement to understand your customer's point of view. The Customer Problem Statement template helps you focus on what matters to create experiences people will love. A well-articulated customer problem statement allows you and your team to find the ideal solution for the challenges your customers face. Throughout the process, you'll also be able to empathize with your customers, which helps you better understand how they perceive your product or service.

<b>I am</b>	<small>Describe customer with 3-4 key characteristics - who are they?</small>	Describe the customer and their attributes here
<b>I'm trying to</b>	<small>List their outcome or "job" the care about - what are they trying to achieve?</small>	List the thing they are trying to achieve here
<b>but</b>	<small>Describe what problems or barriers stand in the way - what bothers them most?</small>	Describe the problems or barriers that get in the way here
<b>because</b>	<small>Enter the "root cause" of why the problem or barrier exists - what needs to be solved?</small>	Describe the reason the problems or barriers exist
<b>which makes me feel</b>	<small>Describe the emotions from the customer's point of view - how does it impact them emotionally?</small>	Describe the emotions the result from experiencing the problems or barriers



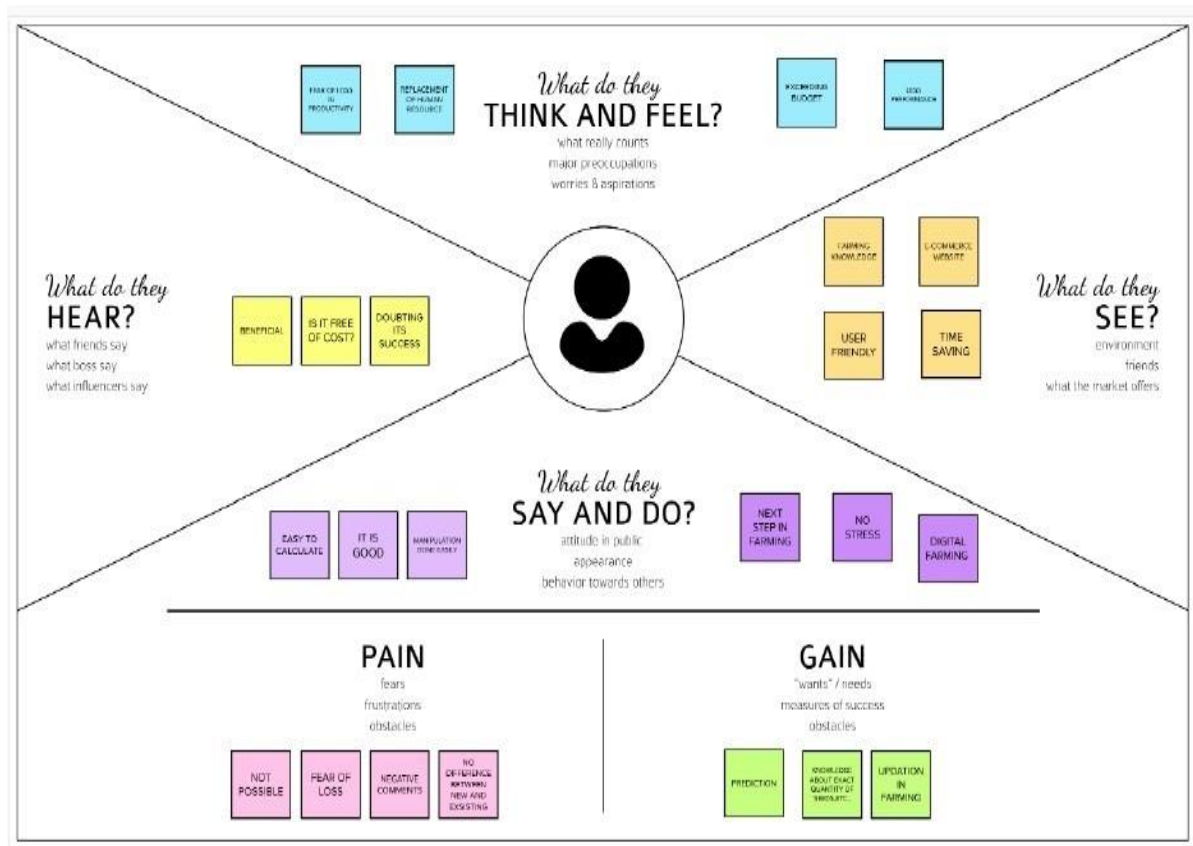
Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	Industrial expert	Know climatic condition	Cannot predict	For marketability&profitabilty	Frustrated
PS-2	Land Owner	Resist crops from pests and diseases	Affected from pest	To avoid farm from pest attack and disease	worried

### 3. IDEATION & PROPOSED SOLUTION

#### 3.1 Empathy Map Canvas

An empathy map is a simple, easy-to-digest visual that captures knowledge about a user's behaviours and attitudes. It is a useful tool to helps teams better understand their users.

Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user's perspective along with his or her goals and challenges.



### 3.2 Ideation & Brainstorming

Brainstorm:

**Harini**



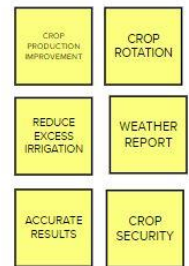
**Chithranjani**



**Padma Priya**



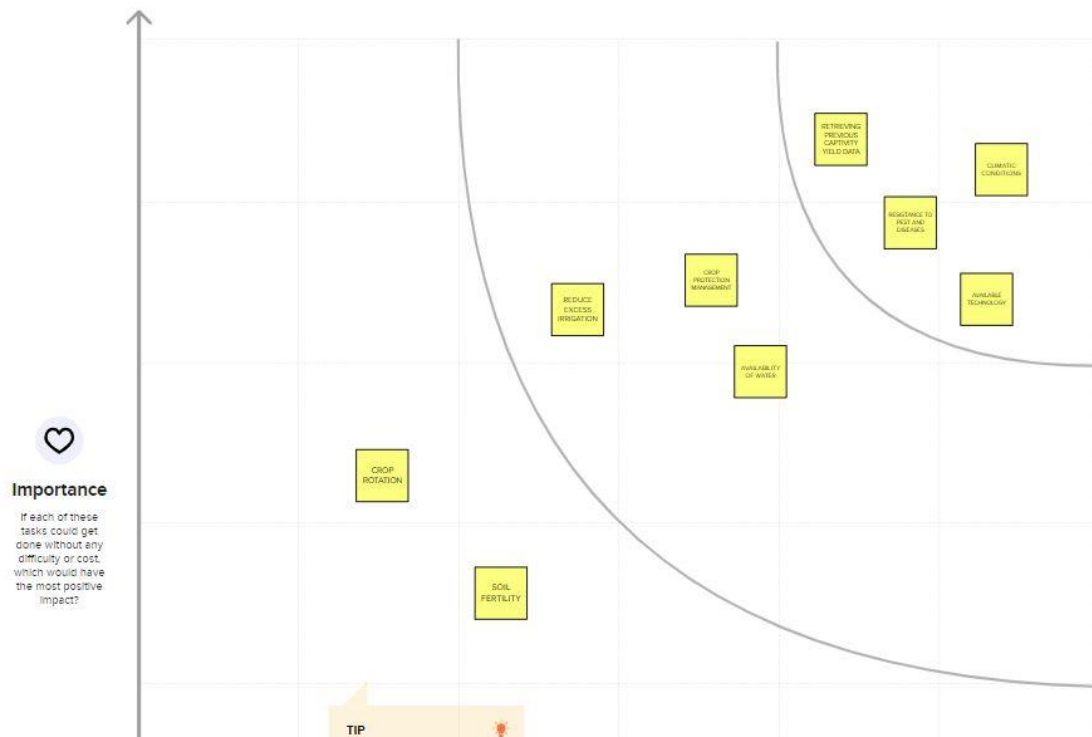
**Siva Gayathri**



Group ideas:



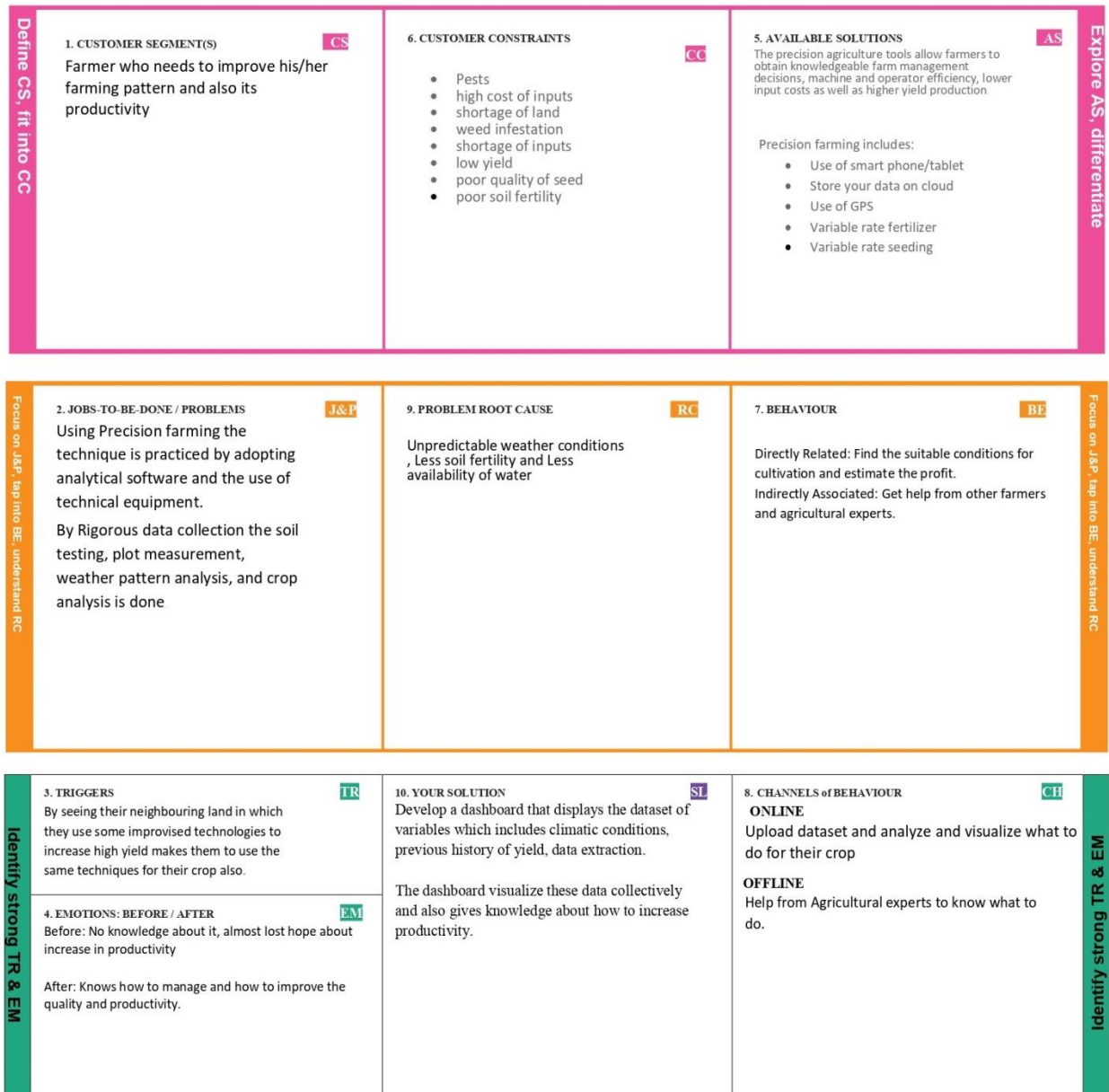
Idea prioritization:



### 3.3 Proposed Solution

S. No	Parameter	Description
1.	Problem Statement (Problem to be solved)	A FARMER who needs to know about the previous year datas (which includes soil fertility, weather conditions& regions) and adapt large farms for utilizing precision farming to increase yields, reduce waste.
2.	Idea / Solution description	This project aims at creating a dashboard which visualise the data like weather report, yield in previous year, retrieving data from previous records, irrigation facility. These datas are all together displayed in dashboard which clearly provide idea about how to gain profit in production yield.
3.	Novelty / Uniqueness	The features of this dashboard is all the data which is needed will be displayed in one which is easily understandable. Not other dashboards are required separately.
4.	Social Impact / Customer Satisfaction	It will be helpful to farmers, Investors, Land owners and Business persons to gain profit in yield.
5.	Business Model (Revenue Model)	E-Commerce websites is the best way to display the dashboard because many people will come across it. And also focusing more on peoples related to farming or dependent on it.
6.	Scalability of the Solution	It has the high database storage so that the data stored can be viewed and retrieved anytime, anywhere.

### 3.4 Problem Solution fit



## 4. REQUIREMENT ANALYSIS

### 4.1 Functional requirement

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through SMS Link
FR-2	User Confirmation	Confirmation via SMS Confirmation via OTP

## 4.2 Non-Functional requirements

Following are the non-functional requirements of the proposed solution.

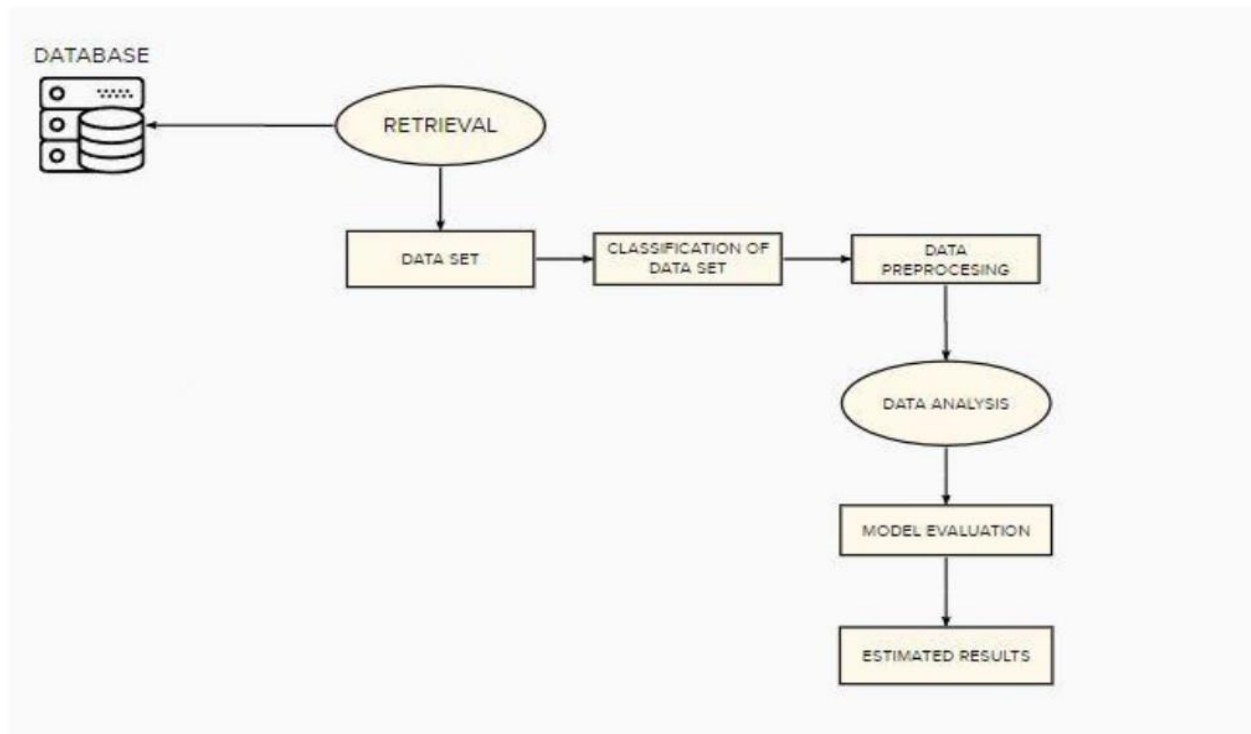
FR No.	Non-Functional Requirement	Description
NFR-1	<b>Usability</b>	The features of this dashboard is, all the data which is needed will be displayed in one which is easily understandable. Not other dashboards are required separately.
NFR-2	<b>Security</b>	Not all the intruders are allowed to access any particular or whole data from dashboard. Only recognized users can access the resource.
NFR-3	<b>Reliability</b>	Either any new Visualization and dashboards are added or erased it won't affect other dashboards. And also the added/erased dashboard will be retrieved soon.
NFR-4	<b>Performance</b>	It will be helpful to farmers, Investors, Land owners and Business persons to gain profit in yield.
NFR-5	<b>Availability</b>	Precision farming can improve time management, reduce water and chemical use, and produce healthier crops and higher yields.
NFR-6	<b>Scalability</b>	It has the high database storage so that the data stored can be viewed and retrieved anytime, anywhere.

## 5. PROJECT DESIGN

### 5.1 Data Flow Diagrams

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



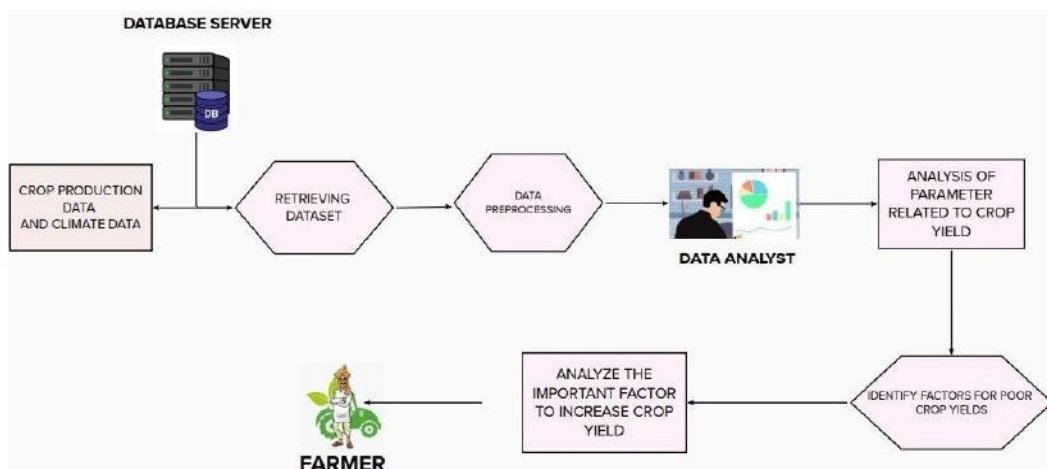


## 5.2 Solution & Technical Architecture

### Solution Architecture:

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

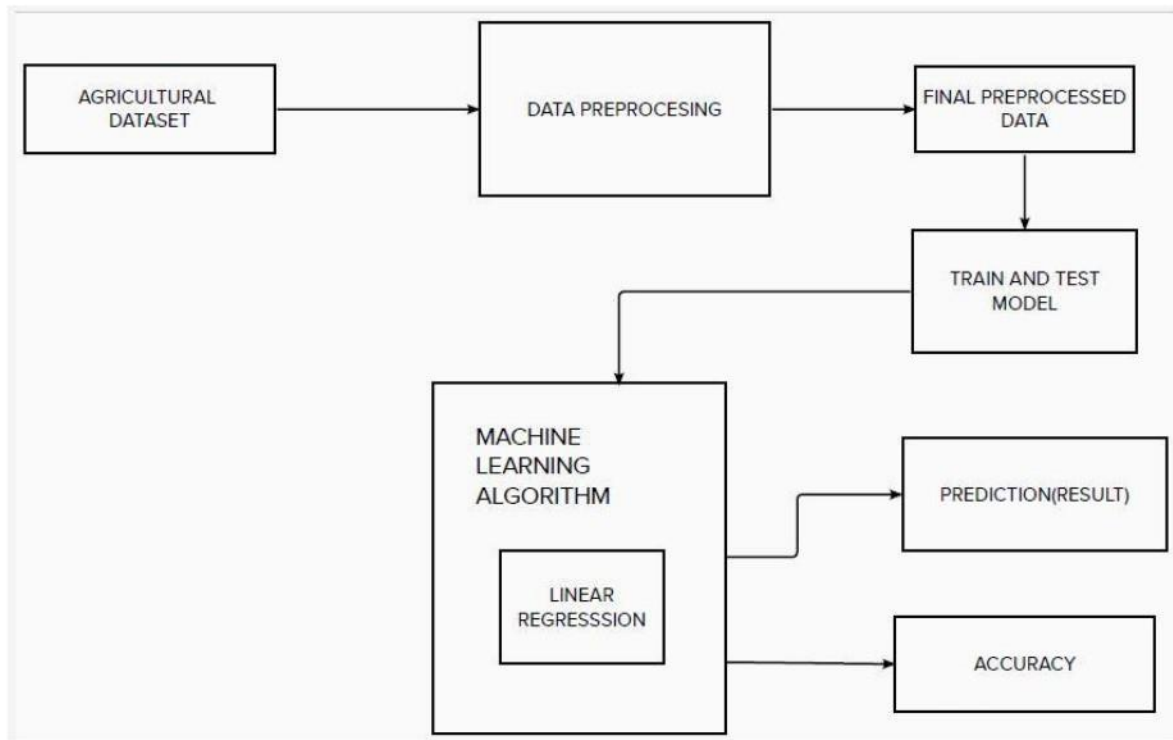
- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed, and delivered.





### Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2



### Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript
2.	Application Logic-1	Login as a user(farmer) in application	Java / Python
3.	Application Logic-2	Login as admin in the application	IBM Watson STT service
4.	Application Logic-3	Login as merchant the application	IBM Watson Assistant
5.	Database	Data related to crop production in previous and also crop data	MySQL, NoSQL
6.	Cloud Database	IBM Watson cloud provides storage	IBM DB2, IBM Cognos
7.	External API-1	Weather APIs are Application Programming Interfaces that allow you to connect to large databases of weather forecast and historical information.	IBM Weather API
8.	External API-2	Soil testing is a quick and accurate method to determine the relative acidity of the soil and the level of several essential nutrients needed.	Soil Test API
9.	Machine Learning Model	It is mostly used for finding out the relationship between variables and forecasting.	Linear Regression
10.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :I1	Local, Cloud Foundry, Kubernetes, etc.

### Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	A software wherein original source code is made freely available and may be redistributed and modified according to the user requirement.	Apache Spark and Hadoop
2.	Security Implementations	crop-related diseases	multispectral camera sensors mounted on drones
3.	Scalable Architecture	A 3-tier architecture wherein application gets data from various sources, manipulates it, stores them in IBM Cloud and visualize them through IBM Cognos	IBM Cloud, IBM Cognos
4.	Availability	The application being developed is made available to all users(farmers).	Cognos Analytics
5.	Performance	Multiple technologies and services that will improve the usability in agricultural activities	Robots, IoT Agriculture sensors.

### 5.3 User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user/Laptop user, etc.,)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application.	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through SMS link.	I can register & access the dashboard with SMS Link	Medium	Sprint-2
		USN-4	As a user, I can register for the application through Gmail.	Through my registration mail also I can login	Low	Sprint-3
	Login	USN-5	As a user, I can log into the application by entering email & password	I can access with my user login id and password	High	Sprint-1
	Dashboard	USN-6	Able to use all the options present in the dashboard.		Medium	Sprint-2
	Contribution	USN-7	With the results obtained from the dashboard, can calculate profit or loss.	Profit or loss	Medium	Sprint-4
Customer (Web user)	Access of resources	USN-8	Accessing with my own login id and password	Only I can access through my user login id and password others shouldn't	High	Sprint -3
Administrator	Manipulation of dataset		Can do modification or updation and storing of data	Modify ,Update,store	High	Sprint-1

## 6. PROJECT PLANNING & SCHEDULING

### 6.1 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user I can register for the application by entering my email ,password ,and confirming my password.	5	High	Padma priya G
Sprint-1		USN-1	As a user I will receive confirmation email once registered for the application	7	Medium	Chithranjani JR
Sprint-1		USN-1	As a user I can register for the application through gmail	4	Medium	Sivagayathri R
Sprint-1	Login	USN-2	As a user/administrator I can login to the application by entering email & password	4	High	Harini Ga
Sprint-2	Dashboard	USN-3	As a user I can use a dashboard to analyse the crop production	8	High	Harini Ga
Sprint-2	Viewproducts	USN-3	As a user I can view the information related to crop production	3	High	Chithranjani JR
Sprint-2	Updating Crop details	USN-4	As a administrator I can enlist the crop Data	5	Medium	Padma priya G

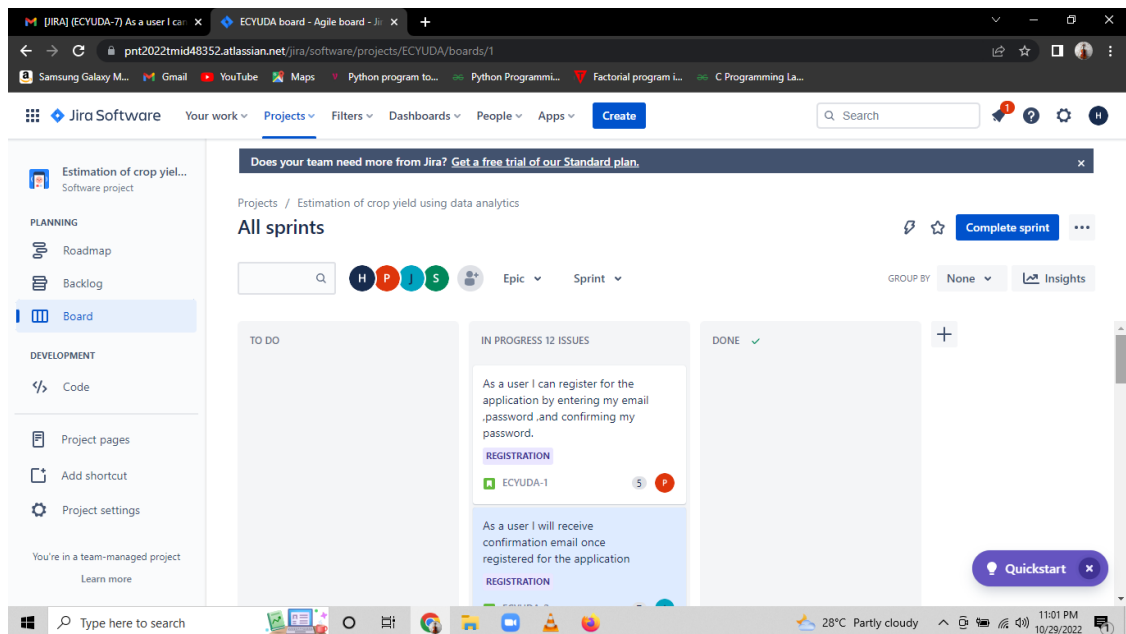
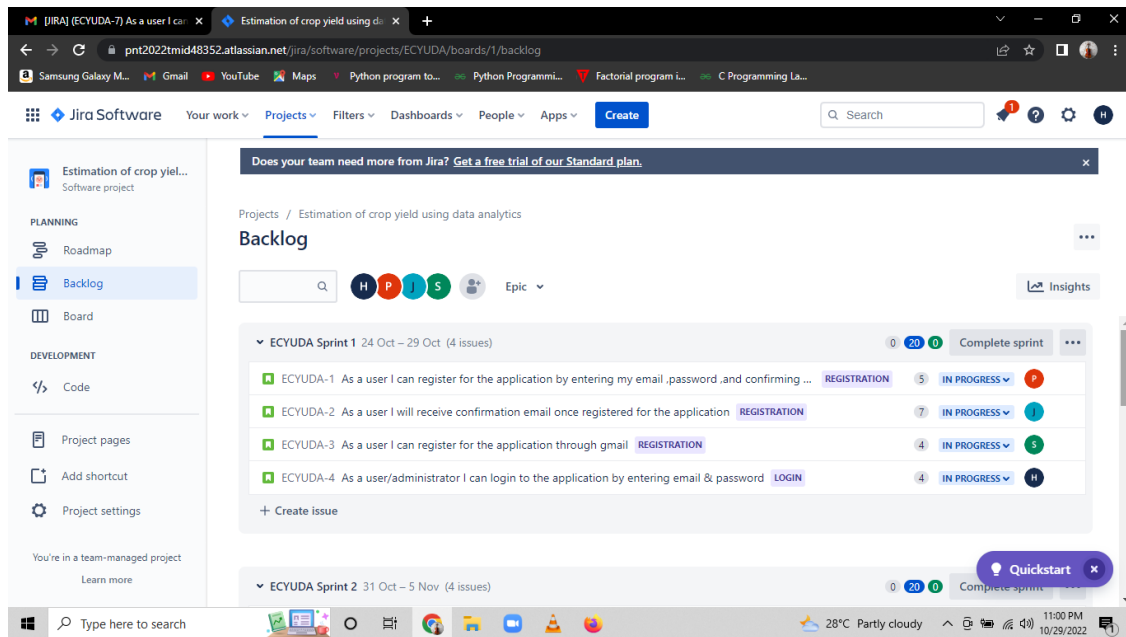
Estimation:

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-2	Crop analysis	USN-3	As a user I can analyse the which factor affect crop production	4	Medium	Sivagayathri R
Sprint-3	Report generation	USN-2	As a user I can create the reports based on customer need	8	High	Sivagayathri R
Sprint-3	Alerts /remainder	USN-2	As a user I can give remainder about which factor does not affect crop production	2	Medium	Chithranjani JR
Sprint-4	Subscription and deployment	USN-5	As a user I can manage the subscription plan details and ensure the application is successfully deployed	5	High	Padma priya G
Sprint-4	Deployment incloud	USN-5	To ensure the application is successfully deployed in cloud	5	High	Harini Ga

### 6.2 Sprint Delivery Schedule

	OCT						NOV						NOV						NOV								
	25	26	27	28	29	30	31	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Sprints	ECYUDA Sprint 1						ECYUDA Sprint 2						ECYUDA Sprint 3						ECYUDA Sprint 4								
> <a href="#">ECYUDA-5 Registration</a>																											
> <a href="#">ECYUDA-6 Login</a>																											
> <a href="#">ECYUDA-11 Dashboard</a>																											
> <a href="#">ECYUDA-12 View products</a>																											
> <a href="#">ECYUDA-13 Updating crop details</a>																											
> <a href="#">ECYUDA-14 Crop Analysis</a>																											
> <a href="#">ECYUDA-17 Reports generation</a>																											
> <a href="#">ECYUDA-18 Alerts/remainder</a>																											
> <a href="#">ECYUDA-20 Subscription and deployment</a>																											
> <a href="#">ECYUDA-21 Deployment in cloud</a>																											

## 6.3 Reports from JIRA



## 7. CODING & SOLUTIONING (Explain the features added in the project along with code)

### 7.1 Feature 1:

#### 1.REGISTRATION

##### Importing sqlite and declaration

from tkinter import \*

import sqlite3

```
root = Tk()
root.geometry('800x500')
root.title("Registration Form")
```

```
Fullname = StringVar()
Email = StringVar()
var = IntVar()
c = StringVar()
var1 = IntVar()
```

### **Defining Database:**

```
def database():
    name1 = Fullname.get()
    email = Email.get()
    gender = var.get()
    region = c.get()
    category = var1.get()
    password = var.get()
```

### **Connecting database:**

```
conn = sqlite3.connect('details.db')
with conn:
    cursor = conn.cursor()
    cursor.execute(
        'CREATE TABLE IF NOT EXISTS USER (Fullname TEXT,Email TEXT,Gender TEXT,region TEXT,category TEXT,password TEXT,confirm password TEXT)')
    cursor.execute('INSERT INTO USER (FullName,Email,Gender,region,category) VALUES(?,?,?,?,'
        (name1, email, gender, region, category))
    conn.commit()
```

### **LOGIN:**

#### **Importing sqlite and function tool:**

```
from tkinter import *
from functools import partial
import sqlite3
```

#### **Validating username and password:**

```
def validateLogin(username, password):
    print("username entered :", username.get())
    print("password entered :", password.get())
    return
```

### **Connecting and defining database:**

```
def database():  
    name = username.get()  
    pswd=password.get()  
  
    conn = sqlite3.connect('login.db')  
    with conn:  
        cursor = conn.cursor()  
        cursor.execute(  
            'CREATE TABLE IF NOT EXISTS VIEWER (name TEXT,password TEXT)')  
        cursor.execute('INSERT INTO VIEWER (name,password) VALUES(?,?)',(name,password))  
        conn.commit()
```

## 7.2 Feature 2

### **2.Creating button and linking it with browser which shows exploration:**

#### **Importing libraries:**

```
from tkinter import *  
from tkinter import messagebox  
from tkinter import ttk  
import tkinter  
import webbrowser
```

#### **Command for linking browser:**

```
def link():  
  
    webbrowser.open_new("file:///C:/Users/dell/Downloads/Estimation%20of%20crop%20yield.htm  
l")  
  
    nut = ttk.Button(root, text='Crops and details',command=link)  
  
    nut.pack()  
    root.mainloop()  
    root.update()
```

### **3.Notificaion or alert message:**

#### **Creating notification message for updation:**

```
import tkinter  
from tkinter import messagebox  
  
# Message Box  
messagebox.showinfo("Notification/Alert", "DETAILS ADDED")
```

## 7.3 Database Schema (if Applicable)

### **Registration :creation of table in sqlite**

```
import sqlite3
```

```

con=sqlite3.connect("details.db")
cursor=con.cursor()
stmt="""CREATE TABLE registration(
    name VARCHAR(20),
    email VARCHAR(20),
    gender VARCHAR(19),
    region VARCHAR(19),
    category VARCHAR(20),
);"""

```

```

con.commit()
cursor.execute(stmt)
con.close()
print("Table Created")

```

### **Login :creation of table in sqlite**

import sqlite3

```

con=sqlite3.connect("login.db")
cursor=con.cursor()
stmt="""CREATE TABLE login(
    name VARCHAR(20),
    pswd VARCHAR(20),
);"""

```

```

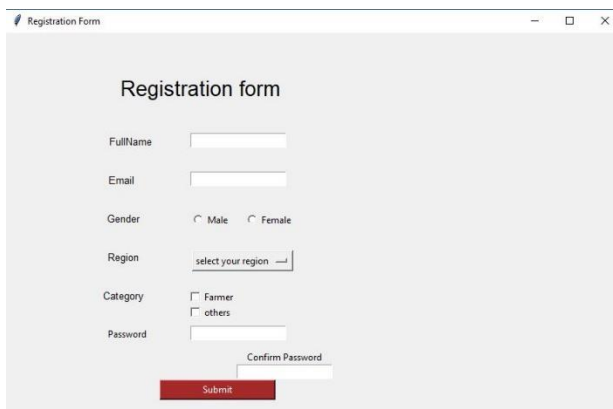
con.commit()
cursor.execute(stmt)
con.close()
print("Table Created")

```

## **8. TESTING**

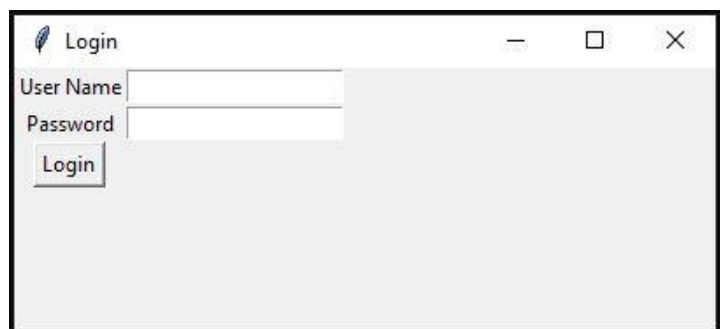
### **8.1 Test Cases**

**REGISTRATION AND LOGIN:**



The screenshot shows a web browser window titled "Registration Form". The form is titled "Registration form" and contains the following fields and controls:

- FullName**: A text input field.
- Email**: A text input field.
- Gender**: Two radio buttons labeled "Male" and "Female".
- Region**: A dropdown menu with the text "select your region".
- Category**: Two checkboxes labeled "Farmer" and "others".
- Password**: A text input field.
- Confirm Password**: A text input field.
- Submit**: A red button at the bottom of the form.

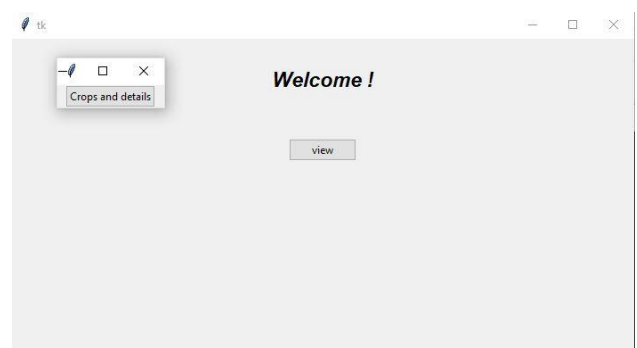
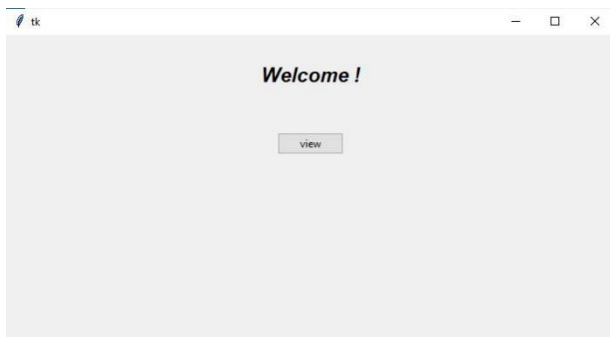


The screenshot shows a web browser window titled "Login". The form is titled "Login" and contains the following fields and controls:

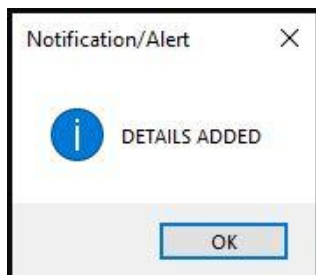
- User Name**: A text input field.
- Password**: A text input field.
- Login**: A button below the password field.



## VIEWING OF EXPLORATION:



## NOTIFICATION /ALERT:



## 8.2 User Acceptance Testing

Resolution	Severity1	Severity2	Severity3	Severity4	Subtotal
By Design	7	3	2	3	15
Duplicate	2	0	2	0	4
External	2	1	0	3	6
Fixed	8	3	5	15	31
Not Reproduced	0	0	0	0	0
Skipped	0	0	2	1	3
Won'tFix	0	4	3	2	9
Totals	19	11	14	24	68

Section	TotalCases	Not Tested	Fail	Pass
PrintEngine	7	0	0	7
ClientApplication	51	0	3	48

Security	2	0	1	1
OutsourceShipping	3	0	1	2
ExceptionReporting	9	0	2	7
FinalReportOutput	4	0	1	3

## **9. RESULTS**

### **9.1 Performance Metrics**

It will be helpful to farmers, Investors, Land owners and Business persons to gain profit in yield.

## **10. ADVANTAGES & DISADVANTAGES**

### **Advantages:**

- Increases the yielding capacity of the crops
- Improves the soil structure
- Crop rotations ensure water conservation
- Pest control and weed management

### **Disadvantages:**

- Difficult to specify
- Not appropriate in certain conditions
- Improper implementation may lead to more harm than good

## **11. CONCLUSION**

As a result of penetration of technology into agriculture field, there is a marginal improvement in the productivity. The innovations have led to new concepts like digital agriculture, smart farming, precision agriculture etc. In the literature, it has been observed that analysis has been done on agriculture soils, hidden patterns discovery using data set related to climatic conditions and crop yields data. The activities of agriculture field are numerous like weather forecasting, soil quality assessment, seeds selection, crop yield prediction etc. In this survey, the specific activity, crop yield prediction has been surveyed and the major trends have been identified.

## **12. FUTURE SCOPE**

We have to collect all required data by giving GPS locations of a land and by taking access from Rain forecasting system of by the government, we can predict crops by just giving GPS location. Also, we can develop the model to avoid over and under crisis of the food.

## **13. APPENDIX**

Source Code

Cognos analytics link:

For dashboard:

[https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my\\_folders%2FState%2Bdashboard&action=view&mode=dashboard&subView=model0000018483df7878\\_00000000](https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FState%2Bdashboard&action=view&mode=dashboard&subView=model0000018483df7878_00000000)

[https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my\\_folders%2FCrop%2Bdashboard&action=view&mode=dashboard&subView=model0000018483c316a7\\_00000000](https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FCrop%2Bdashboard&action=view&mode=dashboard&subView=model0000018483c316a7_00000000)

For data visualization:

[https://us3.ca.analytics.ibm.com/bi/?perspective=explore&pathRef=.my\\_folders%2FCrops%2Band%2Bseasons](https://us3.ca.analytics.ibm.com/bi/?perspective=explore&pathRef=.my_folders%2FCrops%2Band%2Bseasons)

[https://us3.ca.analytics.ibm.com/bi/?perspective=explore&pathRef=.my\\_folders%2FCrop%2Byear%2Band%2Bseason](https://us3.ca.analytics.ibm.com/bi/?perspective=explore&pathRef=.my_folders%2FCrop%2Byear%2Band%2Bseason)

For Report:

[https://us3.ca.analytics.ibm.com/bi/?pathRef=.my\\_folders%2FCrop%2Breport&action=run&format=HTML&prompt=false](https://us3.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FCrop%2Breport&action=run&format=HTML&prompt=false)

Linear Regression:

<file:///C:/Users/dell/Downloads/Estimation%20of%20crop%20yield.html>

GitHub & Project Demo Link

GitHub:

<https://github.com/IBM-EPBL/IBM-Project-38164-1660373931>

Demo Link:

<https://www.mediafire.com/file/gl7yetkkvv59zol/Final-1.mp4/file>