

A PROJECT REPORT ON

**ESTIMATE THE CROP YIELD USING DATA
ANALYTICS**

DOMAIN: DATA ANALYTICS

DONE BY

TEAM ID: PNT2022TMID09610

Team Size : 5

Team Leader : Chandra Prakash S

Team member : Amirthanand S

Team member : Dani J

Team member : Joshua

Team member : Kamalakannan

TABLE OF CONTENTS

1.ABSTRACT

2. INTRODUCTION

1. Project Overview
2. Purpose

3. LITERATURE SURVEY

1. Existing problem
2. References
3. Problem Statement Definition

4. IDEATION & PROPOSED SOLUTION

1. Empathy Map Canvas
2. Ideation & Brainstorming
3. Proposed Solution
4. Problem Solution fit

5. REQUIREMENT ANALYSIS

1. Functional requirement
2. Non-Functional requirements

6. PROJECT DESIGN

1. Data Flow Diagrams
2. Solution & Technical Architecture
3. User Stories

7. PROJECT PLANNING & SCHEDULING

1. Sprint Planning & Estimation
2. Sprint Delivery Schedule

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

1. Feature 1
2. Feature 2

8.3 Database Schema (if Applicable)

9. TESTING

1. Test Cases
2. User Acceptance Testing

10. RESULTS

1. Performance Metrics

11. ADVANTAGES & DISADVANTAGES

12. CONCLUSION

13. FUTURE SCOPE

14. APPENDIX

GitHub & Project Demo Link

1.ABSTRACT

One of the most important occupations for human survival is Agriculture and especially the majority of the population in India is into this. Due to variations in climatic conditions and various other challenges, the Agrarian sector in India is facing rigorous problems to maximize crop productivity. It has become an arduous task to achieve the desired targets in crop yield. Various factors like monsoon rains, rodents, water scarcity are to be considered, which do have a direct impact on the production and productivity of the crops. And thus, crop yield prediction with recent advancements in technology can be one of the important factors in agriculture practices as the use of technology in agriculture has comparatively increased in the recent years. The present study gives insights on data analytics methods and various machine learning techniques applied to find the crop yield prediction

2.INTRODUCTION

1. Project overview

- ✓ It has become a challenging task to achieve desired targets in Agriculture based crop yield. Various factors 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 years and data analytics is one such trend that has penetrated into the
- ✓ agriculture field. To reach desired crop yield goals has become a difficult undertaking in agriculture.
- ✓ Numerous elements that directly affect the yield and productivity of the crops must be taken into account. One of the crucial aspects of agricultural techniques is the forecast of crop production. Before planting seeds in their fields, farmers require knowledge about crop yield in order to increase agricultural output.
- ✓ In recent years, the use of technology in agriculture has increased, and one such development is the use of data analytics.
- ✓ Thus, a project that would suit the needs of a farmer and at least help them over a borderline to understand and predict or estimate the crop yield was the main aim and was brought to life.

2. Purpose

- ✓ Estimation of Crop yield has become the need of the hour and one easy tool/method that can be used is Data Analytics. The term "data analytics" describes the methods used to analyse data in order to increase productivity and financial gain. In order to examine different behavioural patterns, data is extracted from a variety of sources, cleaned up, and classified. The methods and resources employed change depending on the group or person.
- ✓ The purpose behind this project is to understand the variation in crop yield due to various parameters that can be natural or non-natural

3.LITERATURE SURVEY



1. Existing problem

With the changing of climate, agriculture faces increasing problems with extreme weather events leading to considerable yield losses of crops. Most often, crop plants are sensitive to stresses since they were mostly selected for high yield, and not for stress tolerance. The four most important factors that influence crop yield are soil fertility, availability of water, climate, and diseases or pests.

With such varying parameters, to understand or estimate the patterns with no technological involvements is very difficult. Thus, a solution that is technological and cater to the alterations and provide the predicted solution in a form that can be easily understood by end customers is essential.

2. References

- ✓ [How data analytics is transforming agriculture - ScienceDirect](https://doi.org/10.1016/j.bushor.2017.09.011) - <https://doi.org/10.1016/j.bushor.2017.09.011>
- ✓ [https://www.researchgate.net/publication/329467349_Agriculture Data Analytics in Crop Yield Estimation A Critical Review](https://www.researchgate.net/publication/329467349_Agriculture_Data_Analytics_in_Crop_Yield_Estimation_A_Critical_Review)
- ✓ N. Chergui, M. -T. Kechadi and M. McDonnell, "The Impact of Data Analytics in Digital Agriculture: A Review," 2020 International Multi-Conference on: "Organization of Knowledge and Advanced Technologies" (OCTA), 2020, pp. 1-13, doi: 10.1109/OCTA49274.2020.9151851.
- ✓ D. Elavarasan and P. M. D. Vincent, "Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian

3.3 Problem statement definition

The following instances define the problem of notice. Ram is a farmer who needs a way to understand and predict climatic conditions because he can decide on the safety measures to be followed with regards to the field setup.

Raj is a farmer who needs a way to decide what to grow and

when to grow because he is uncertain of his environmental conditions. Ranil is a grocer and crop distributor who needs to know the overall crop yield turnover because he has to understand his monetary turnover for the year.

Thus, a solution that can cater to all the needs put forth is being formulated.

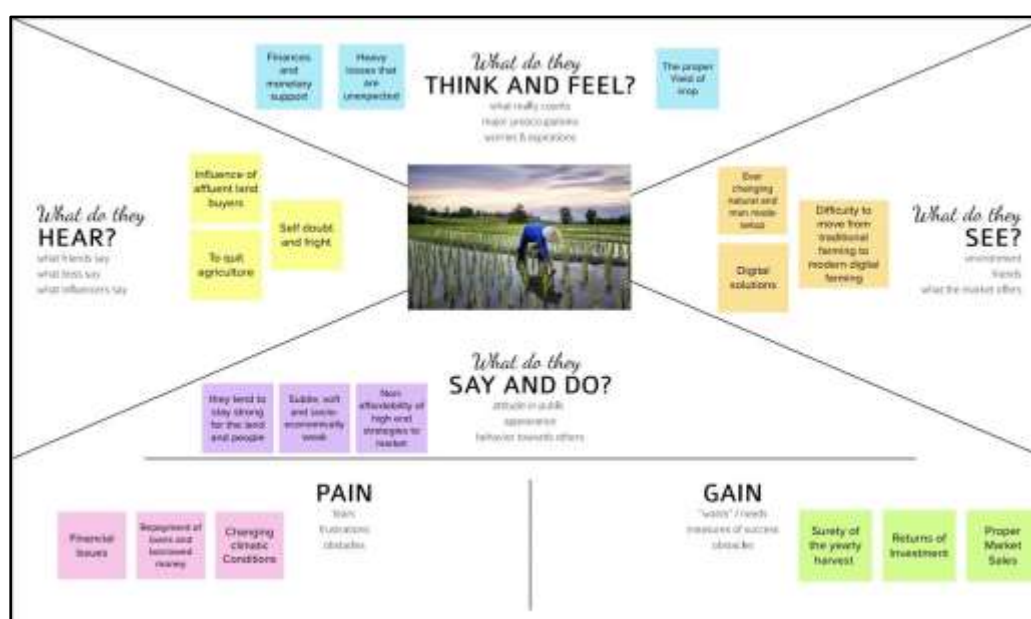
4.IDEATION&PROPOSED SOLUTION

1. EMPATHY MAP CANVAS

A team of four members sat together to discuss on and empathize about the problem that people have been facing with regards to understanding and predicting the yield of crops.

As a part of what the customers or target audience felt, a conclusion was made such that they were concerned about elements like finances, monetary support, heavy unexpected losses, proper yield of crop and certain other unpredictable factors.

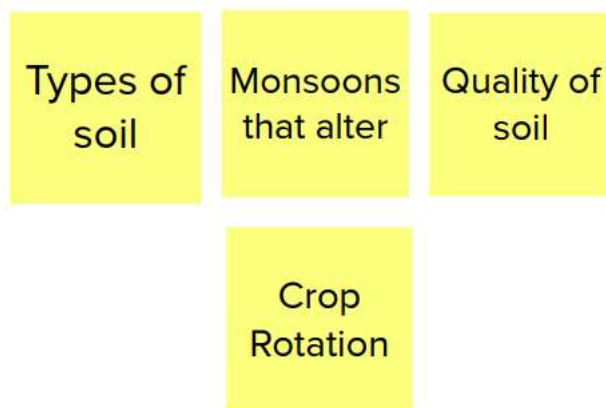
Under the concepts of what they see and on the basis of environment, friends and what the market offers, digital solutions for ever changing natural setup with the difficulty of moving from traditional to modern farming was sought. The specifics of pain and gain along with speculations of the influence of affluent land buyers, self-doubt and fright added with this comes the fear to quit agriculture under the section of what they hear.



2. Ideation & Brainstorming

Ideation and the process of brainstorming was done by initially tracing the problem and defining it. This was followed by pushing in individual ideas about the problem and then grouping it in together under common grounds and making a graph out of the priority provided.

Chandraprakash



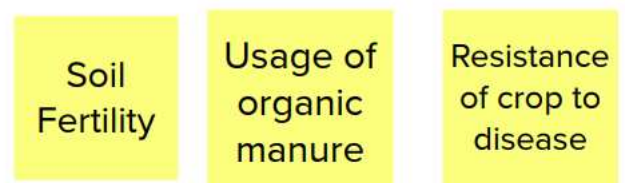
Dani



Amirthananth



Kamalakannan



Joshua



GROUPING:

CROP PARAMETERS

Crop Type

Varied Crop
Growth

Crop Growth
Rate

Interval
between
crops

Crop
Diseases

Crop
Rotation

Resistance
of crop to
disease

SOIL PARAMETERS

Types of soil

Soil Fertility

Quality of
soil

ENVIRONMENTAL SETUP

Environmental
Condition

Monsoons
that alter

Irrigation
type

OTHERS

Yield
prediction

Usage of
organic
manure

Pests and
Control

PRIORITIZATION:



3 Proposed Solution

S.NO.	PARAMETER	DESCRIPTION
1.	Problem Statement (Problem to be solved)	A farmer should predict climatic conditions, decide what to grow & when to grow, should know the overall crop yield turnover and must be able to be sure of the crop yield inspite of the environmental and other parameters
2.	Idea / Solution description	Analysis of important visualization using the previous years' data, creating a dashboard and by going the datasets to obtain most of the insights of Crop production in India.
3.	Novelty / Uniqueness	A one-stop solution for understanding and to get an insight about the previous years' data related to the harvest and cultivation. There is no other setup that's required to
4.	Social Impact / Customer Satisfaction	Availability to all the farmers who need help and as this is a simple approach, understanding issues will not arise.
5.	Business Model (Revenue Model)	A profit can be made by promoting the solution as an easily available mobile application for anyone to access and benefit out of it. Venture joints with government can be made to pull out monetary Benefits.

4.4 Problem Solution fit

Define CS, fit into CC	1. CUSTOMER SEGMENT(S) CS <ul style="list-style-type: none"> Farmers Individuals associated with agricultural activities, cultivation, harvest and sales of the harvested goods. 	6. CUSTOMER CONSTRAINTS CC <p>What constraints prevent your customers from taking action or limit their options of solutions? (i.e. spending power, budget, no cash, network connections, available services)</p> <ul style="list-style-type: none"> Monetary issues Network Issues Lack of Awareness Quality of soil, manure, water etc. 	5. AVAILABLE SOLUTIONS AS <p>Which solutions are available to the customers when they face the problem?</p> <p>so need to get the job done? What have they tried in the past? What pros & cons do these solutions have? (i.e. past and present to an alternative to digital consulting)</p> <ul style="list-style-type: none"> Traditional ways of prediction Precision farming 	Explore AS, differentiate
	2. JOBS-TO-BE-DONE / PROBLEMS J&P <ul style="list-style-type: none"> Help them understand the usage of prediction and software for better results in agriculture Data is to be collected and awareness should be brought in order to orchestrate the above mentioned 	9. PROBLEM ROOT CAUSE RC <ul style="list-style-type: none"> Weather conditions Soil Conditions Water availability Unpredictable weather conditions Pest issues Manure and other usages Crop resistance 	7. BEHAVIOUR BE <ul style="list-style-type: none"> Try to get help from agricultural experts Try to take up non-natural means of cultivation for quicker harvest 	

Identify strong TR & EM	3. TRIGGERS TR <p>What triggers customers to act? (i.e. seeing their neighbor planting other plants, reading about a new efficient solution in the news)</p> <p>Seeing their neighbor farmers have a better yield by usage of natural or non-natural means</p>	10. YOUR SOLUTION SL <p>If you are working on an existing business, write down your current solution that fits to this canvas, and think how much it fits (scale)</p> <p>If you are working on a new business proposition, then think if there isn't your fit in the current solution up with a solution that fits within customer readiness, where a problem your solution addresses effectively</p> <p>An interactive, visualization dashboard that precisely indicates the predictions of the needed parameters for farmers to get insights about what is needed. It is a one-stop solution and there is no extra setup that is required.</p>	8. CHANNELS of BEHAVIOUR CH <p>8.1 ONLINE: What kind of actions do customers take online? (browse online channels from 47)</p> <p>None</p> <p>8.2 OFFLINE: What kind of actions do customers take offline? (attend offline channels from 47 and use them for customer development)</p> <ul style="list-style-type: none"> Trying to use pesticides and fertilizers that increase gain but cause harm Irrigation channel changes 	Identify strong TR & EM
	4. EMOTIONS: BEFORE / AFTER EM <p>How do customers feel when they face a problem or a job and afterwards? (e.g. being insecure & confident, in control / out of control, communication strategy & design)</p> <p>Before: insecure > strength / trust of growths After</p>			

4. REQUIREMENT ANALYSIS

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 Signup	Registration through Gmail Registration through IBM
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	Data Collection	
FR-4	Data Processing	Data cleaning, removal of noise and obsolete data
FR-5	Visualization Tool	Graphical visualization choices

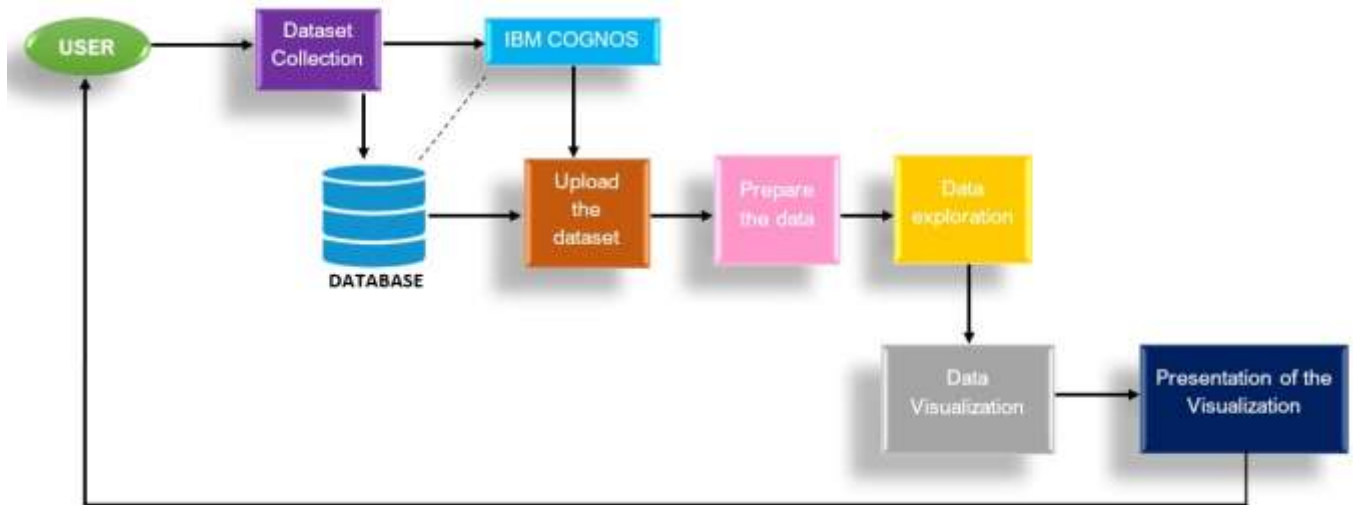
2.NON-FUNCTIONAL REQUIREMENTS

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

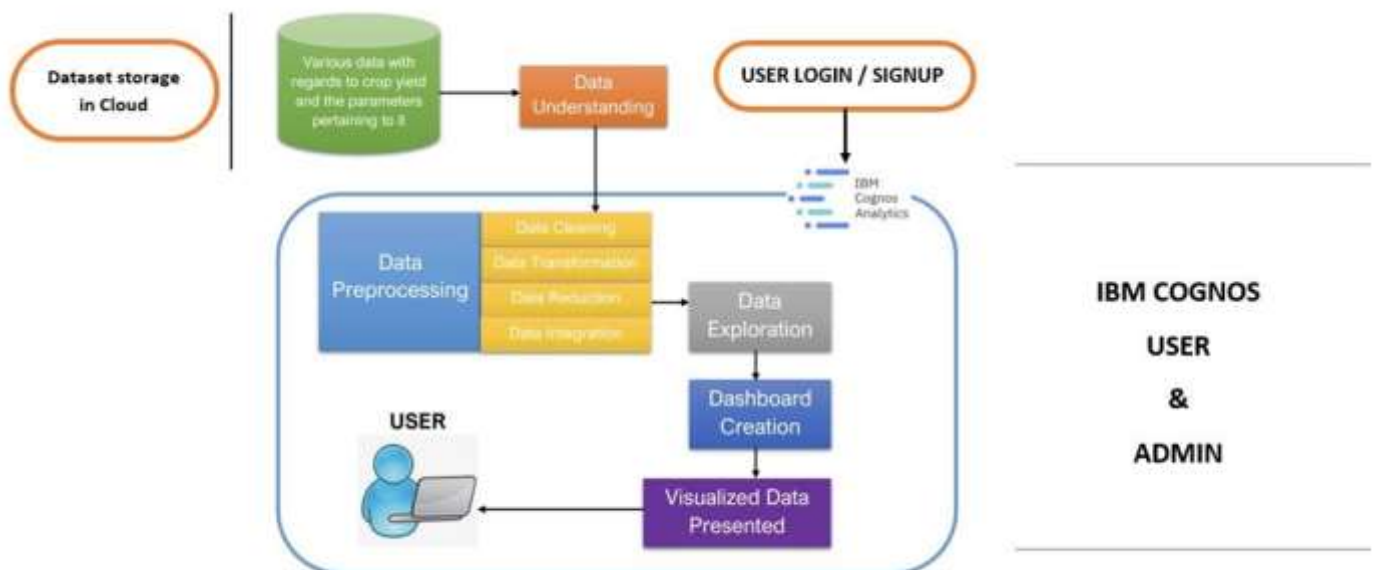
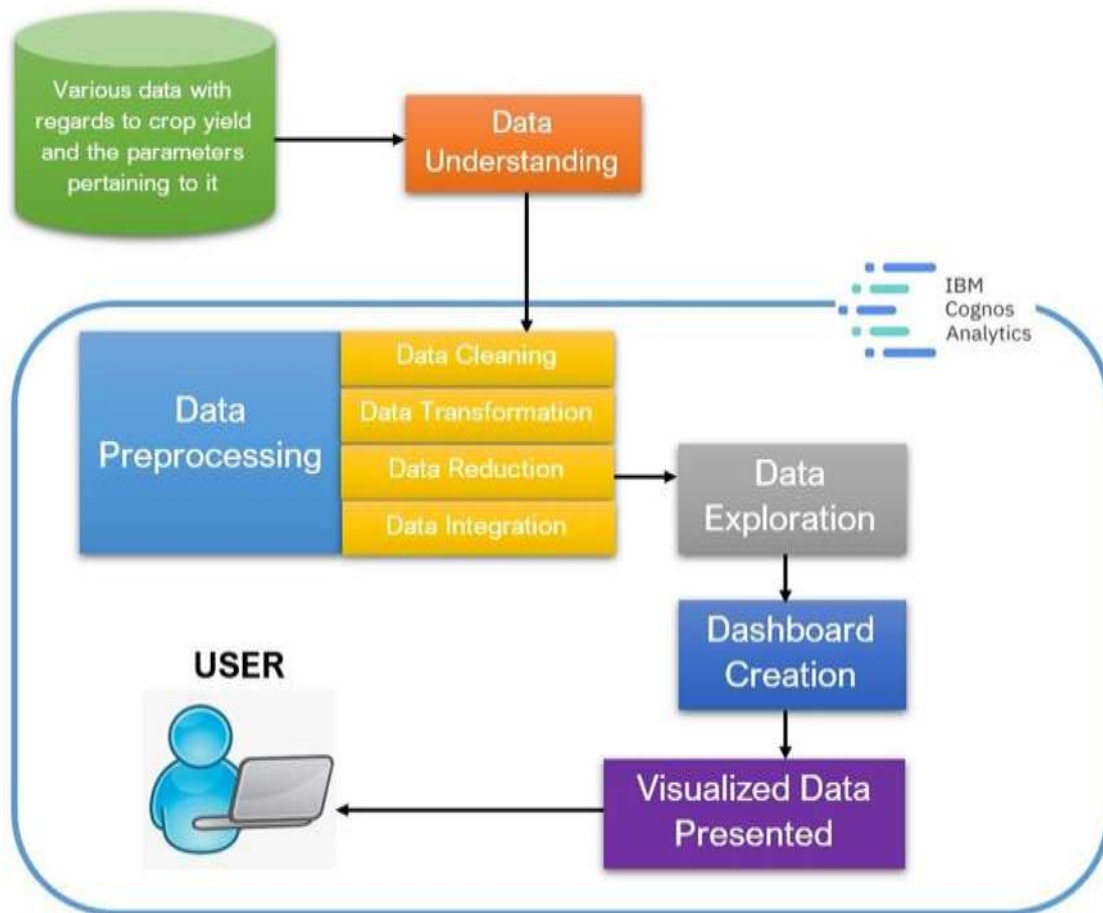
FR No.	Non-Functional Requirement	Description
NFR-1	Usability	Ease of usage along with ease in-access of tools and features
NFR-2	Security	Access to resources through two factor authentication and credentials
NFR-3	Reliability	There should be no crashes or loss of data or processes
NFR-4	Performance	High speed rendering of visualization and other readily available features
NFR-5	Availability	Should be available on demand
NFR-6	Scalability	Should be able to incorporate as many visualizations and datasets as possible

6.PROJECT DESIGN

6.1 Data Flow Diagram



6.2 Solution & Technical Architecture



USER STORIES

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	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 Google	I can register & access the dashboard with Google Login	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail		Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard	USN-6	As a user, I can freely use my dashboard and explore the features		High	Sprint-1
	Access of Resources	USN-7	As a user, I can use the credentials to access the resources of my application	I can securely access my resources	High	Sprint-2
Administrator	Control over the application	USN-8	I can control the users of the application		High	Sprint-2
Customer	Tools	USN-9	I can perform the required tasks on the application		High	Sprint-1

7. PROJECT PLANNING & SCHEDULING

1. SPRINT PLANNING & ESTIMATION

Milestone Name	Milestone Number	Description	Milestone Type	Team Members
Registration and Login	MS-1	Registration for the application by entering my email, password, and confirming my password.	Mandatory	Chan Dani
		Confirmation email once I have registered for the application	Mandatory	Chan Kamal
		Registration for the application through Google	Optional	Amirth Chan
		Registration for the application through Gmail	Optional	Josh Chan
		Log into the application by entering email & password	Mandatory	Chan Kamal
		Usage of dashboard and exploration of the features	Mandatory	Chan Josh
Dashboard Usage	MS-2	Usage of the credentials to access the resources of my application	Mandatory	Dani Amirth
		Performance of Data manipulations on the application	Mandatory	Kamal Chan
Creation of Deliverables	MS-3	Creation of dashboards with particular datasets	Mandatory	Dani Josh
	MS-4	Predictive analysis can be done	Mandatory	Chan Amirth
	MS-5	Creation of stories with particular datasets	Mandatory	Kamal Amirth
	MS-6	Deliver and export reports according to the dashboards and stories created	Mandatory	Josh Chan

2. SPRINT DELIVERY SCHEDULE

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.	2	High	Chan Dani
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	1	High	Chan Kamal
Sprint-2		USN-3	As a user, I can register for the application through Google	2	Low	Amirth Chan
Sprint-1		USN-4	As a user, I can register for the application through Gmail	2	Low	Josh Chan
Sprint-1	Login	USN-5	As a user, I can log into the application by entering email & password	1	High	Chan Kamal
Sprint-3	Dashboard	USN-6	As a user, I can freely use my dashboard and explore the features	2	High	Chan Josh
Sprint-2		USN-7	As a user, I can use the credentials to access the resources of my application	2	High	Dani Amirth
Sprint-3		USN-8	Performance of Data manipulations on the application	1	High	Kamal Chan
Sprint-3	Visualizations	USN-9	I can create dashboards with particular datasets	2	High	Dani Josh
Sprint-4		USN-10	Predictive analysis can be done	1	High	Chan Amirth
Sprint-3		USN-11	I can create stories with particular datasets	2	High	Kamal Amirth
Sprint-4		USN-12	I can deliver and export reports according to the dashboards and stories created	2	High	Josh Chan

3 BURNDOWN CHART

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

Velocity:

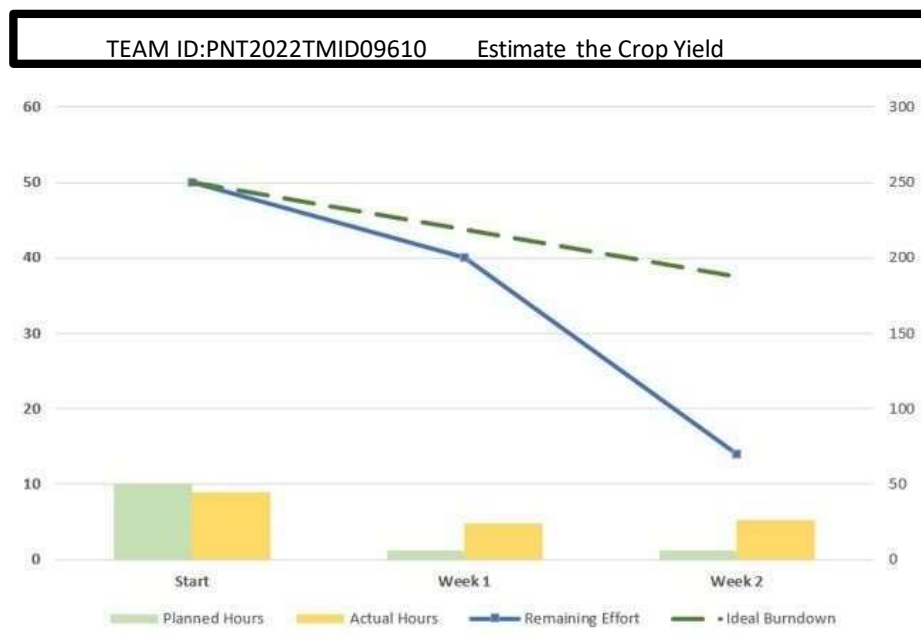
Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

$$AV = \frac{\text{sprint duration}}{\text{velocity}} = \frac{20}{10} = 2$$

$$AV = \text{Sprint Duration} / \text{Velocity} = 20 / 6 = 3$$

Burndown Chart:

A burn down chart is a graphical representation of work left to do versus time. It is often used in agile [software development](#) methodologies such as [Scrum](#). However, burn down charts can be applied to any project containing measurable progress over time.

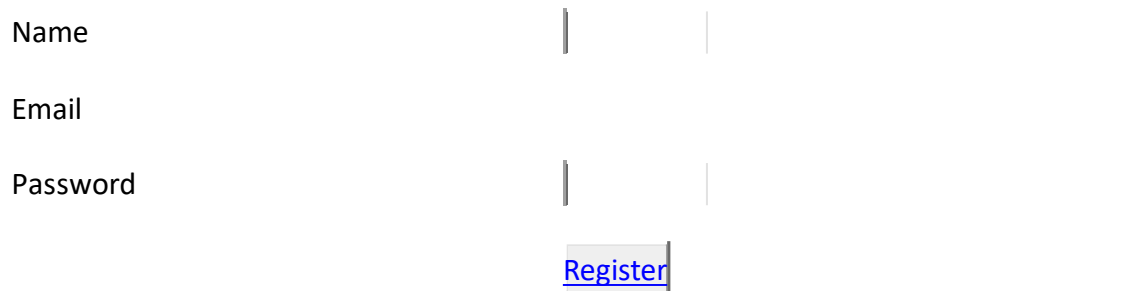


8. CODING & SOLUTIONING

(Explain the features added in the project along with code)

8.1 Feature 1

Registration Form



A registration form with a light gray background. It contains three input fields for 'Name', 'Email', and 'Password', each with a vertical line on the left and a vertical line on the right. Below the 'Password' field is a blue 'Register' button with a white border.

Login Form



A login form with a light gray background. It contains two input fields for 'Email' and 'Password', each with a vertical line on the left and a vertical line on the right. Below the 'Password' field is a gray button.

DATABASE CONNECTIVITY:CODING

CSS COMPONENT

```
mat-card {  
  max-width: 600px;  
  margin: 2em auto;  
  text-align: center;  
  max-height: 600px;
```

```
}
```

```
.header{  
  text-align: center;
```

```
}
```

```
.full-width {  
  width: 80%;  
}
```

```
.button-row {  
  padding-top: 5px;  
}
```

```
.button-row a {  
  margin-right: 8px;  
  text-align: center;  
}
```

```
.forget-password{  
  padding-left: 0px;  
}
```

```
.emailInput{  
  padding-top: 10px;  
}
```

```
.contentBody {  
  padding: 60px 1rem;  
  background :#006064;  
  display: block;  
}
```

```
.aLink{  
  float: right;  
  padding-right: 60px;  
  text-decoration: none;  
}
```

HTML COMPONENT:

```
<mat-card>  
  <mat-card-content>  
    <div class="header">  
      <P>Sign Into Your Account </P>  
    </div>  
    <form (ngSubmit)="onLogin()" name="loginForm" [formGroup]="loginForm">  
      <div class="emailInput">  
        <mat-form-field class="full-width" appearance="outline">  
          <mat-label>Email</mat-label>  
          <input  
            FormControlName="email"  
            matInput  
            placeholder="Enter email address" required  
          />  
          <mat-error *ngIf="!loginForm.controls['email'].valid">
```

```

        Email is required

    </mat-error>

</mat-form-field>

</div>

<div>

    <span>

        <a class="text-link" class="aLink" routerLink="/auth/forgot-
password">Forgot Password?</a>

    </span>

    <mat-form-field class="full-width" appearance="outline">

        <mat-label>Password</mat-label>

        <input formControlName="password" matInput [type]=" hide ?
'password' : 'text'" required />

        <button mat-icon-button matSuffix (click)="hide = !hide"
[attr.aria-label]=" 'Hide Password'"

            [attr.aria-pressed]="hide">

                <mat-icon>

                    {{hide ? 'visibility_off' : 'visibility'}}

                </mat-icon>

            </button>

        <mat-error *ngIf="!loginForm.controls['password'].valid">

            Password is required

        </mat-error>

    </mat-form-field>

</div>

<button mat-flat-button color="primary">Login</button>

</form>

<div class="button-row">

    <p>Create New Account</p>

</div>

```



```
</mat-card-content>

</mat-card>
```

.SPEC.TS COMPONENT:

```
import { ComponentFixture, TestBed } from '@angular/core/testing';

import { LoginComponent } from './login.component';

describe('LoginComponent', () => {
  let component: LoginComponent;
  let fixture: ComponentFixture<LoginComponent>;

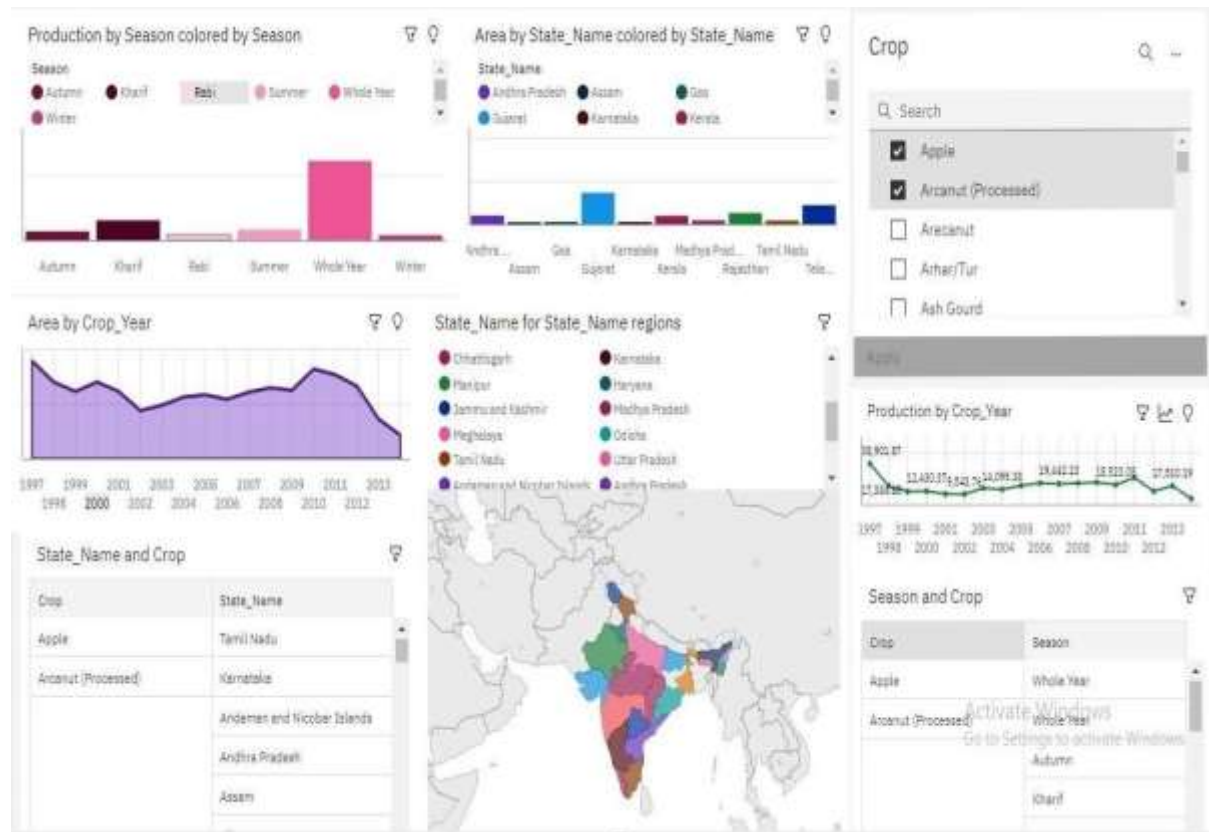
  beforeEach(async () => {
    await TestBed.configureTestingModule({
      declarations: [ LoginComponent ]
    })
    .compileComponents();

    fixture = TestBed.createComponent(LoginComponent);
    component = fixture.componentInstance;
    fixture.detectChanges();
  });

  it('should create', () => {
    expect(component).toBeTruthy();
  });
});
```

8.2 Feature 2

CREATING DASHBOARD USING IBM COGNOS



8.3 Database Schema (if Applicable)

	A	B	C	D	E	F	G	H	I	J
1	State_Na	District_N	Crop_Year	Season	Crop	Area	Production			
2	Andaman	NICOBARS	2000	Kharif	Arecanut	1254	2000			
3	Andaman	NICOBARS	2000	Kharif	Other Kha	2	1			
4	Andaman	NICOBARS	2000	Kharif	Rice	102	321			
5	Andaman	NICOBARS	2000	Whole Ye	Banana	176	641			
6	Andaman	NICOBARS	2000	Whole Ye	Cashewnu	720	165			
7	Andaman	NICOBARS	2000	Whole Ye	Coconut	18168	65100000			
8	Andaman	NICOBARS	2000	Whole Ye	Dry ginger	36	100			
9	Andaman	NICOBARS	2000	Whole Ye	Sugarcane	1	2			
10	Andaman	NICOBARS	2000	Whole Ye	Sweet pot	5	15			
11	Andaman	NICOBARS	2000	Whole Ye	Tapioca	40	169			
12	Andaman	NICOBARS	2001	Kharif	Arecanut	1254	2061			
13	Andaman	NICOBARS	2001	Kharif	Other Kha	2	1			
14	Andaman	NICOBARS	2001	Kharif	Rice	83	300			
15	Andaman	NICOBARS	2001	Whole Ye	Cashewnu	719	192			
16	Andaman	NICOBARS	2001	Whole Ye	Coconut	18190	64430000			
17	Andaman	NICOBARS	2001	Whole Ye	Dry ginger	46	100			
18	Andaman	NICOBARS	2001	Whole Ye	Sugarcane	1	1			
19	Andaman	NICOBARS	2001	Whole Ye	Sweet pot	11	33			
20	Andaman	NICOBARS	2002	Kharif	Rice	189.2	510.84			
21	Andaman	NICOBARS	2002	Whole Ye	Arecanut	1258	2083			
22	Andaman	NICOBARS	2002	Whole Ye	Banana	213	1278			
23	Andaman	NICOBARS	2002	Whole Ye	Black pep	63	13.5			

9. TESTING

9.1 Test Cases

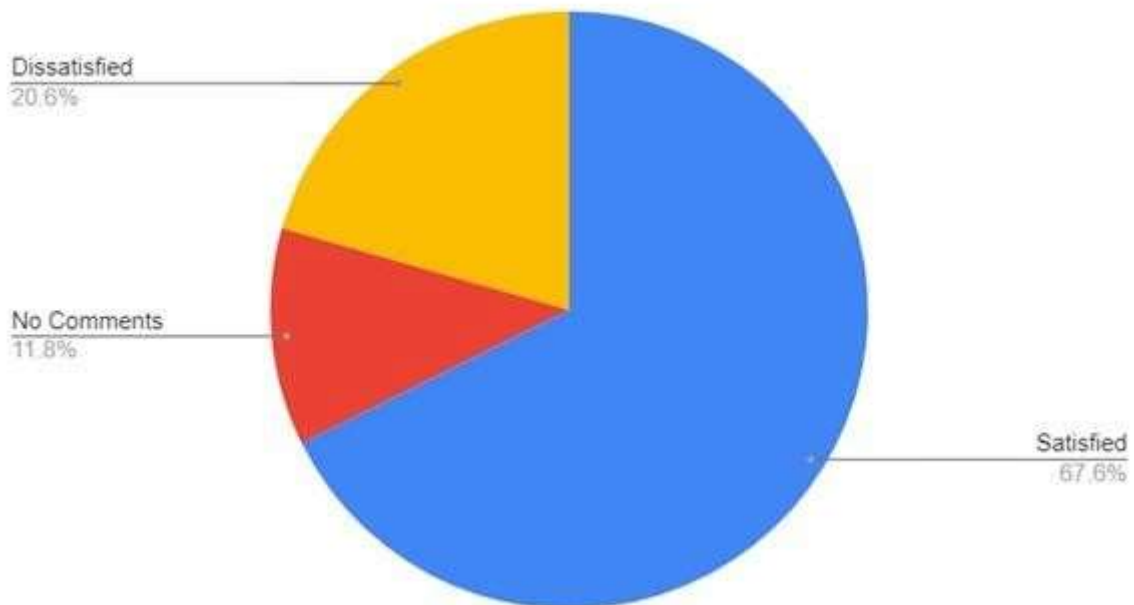
Section	Total Cases	Not Tested	Fail	Pass
Overall process	7	0	0	7
Security	51	0	0	51
Client side	2	0	0	2
Outsource shipping	3	0	0	3
Exception Reporting	9	0	0	9
Final Report Output	4	0	0	4
Version Control	2	0	0	2

Resolution	Severity 1	Severity 2	Severity 3	Severity 4
Use-Easiness	7	5	4	4
Reliability	8	5	9	4
Security	4	5	7	6
Availability	5	7	6	2
Performance	5	4	5	8
Scalability	2	4	3	3

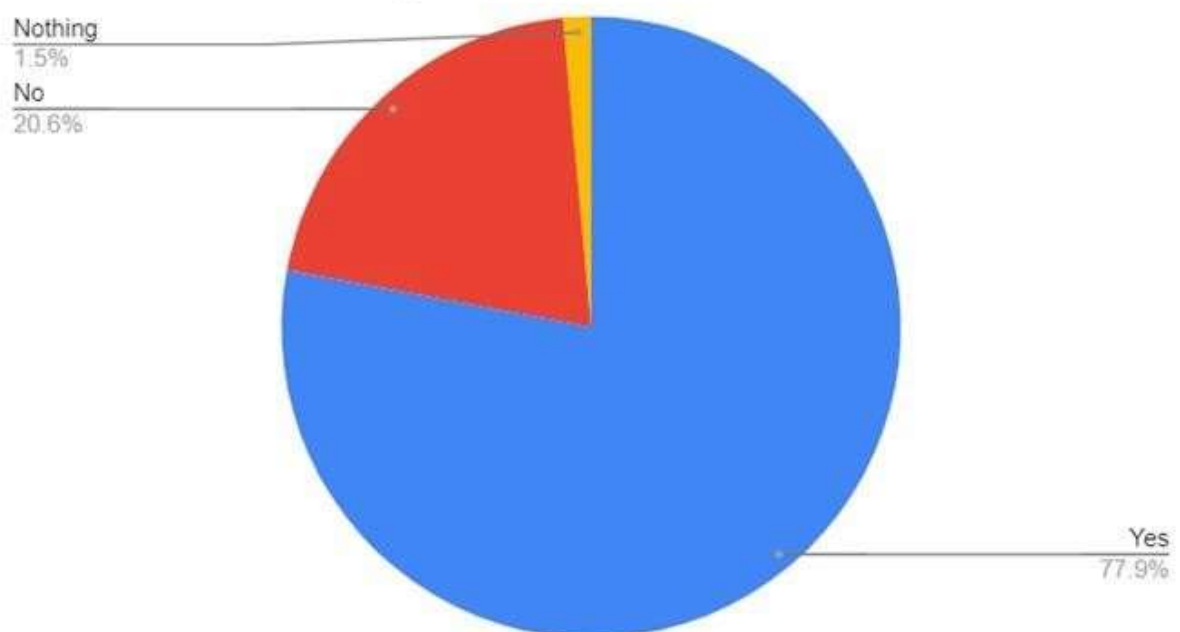
9.2 User Acceptance Testing

This is a survey result on our project's Scalability, User friendliness and Security.

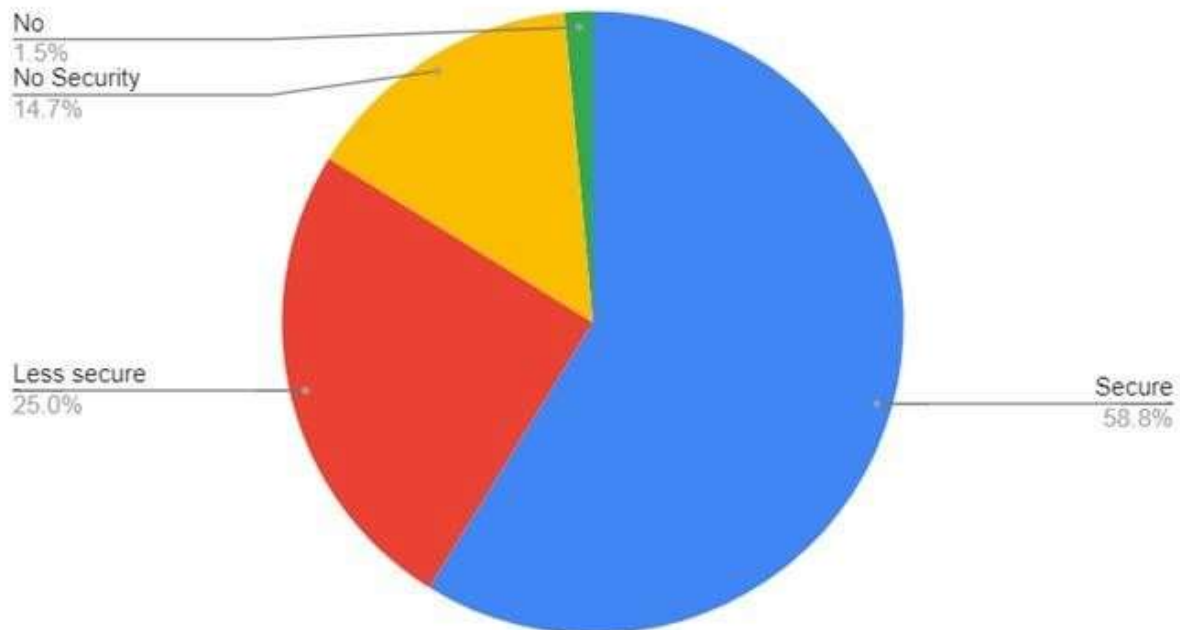
Count of Scalability



Count of User Friendly

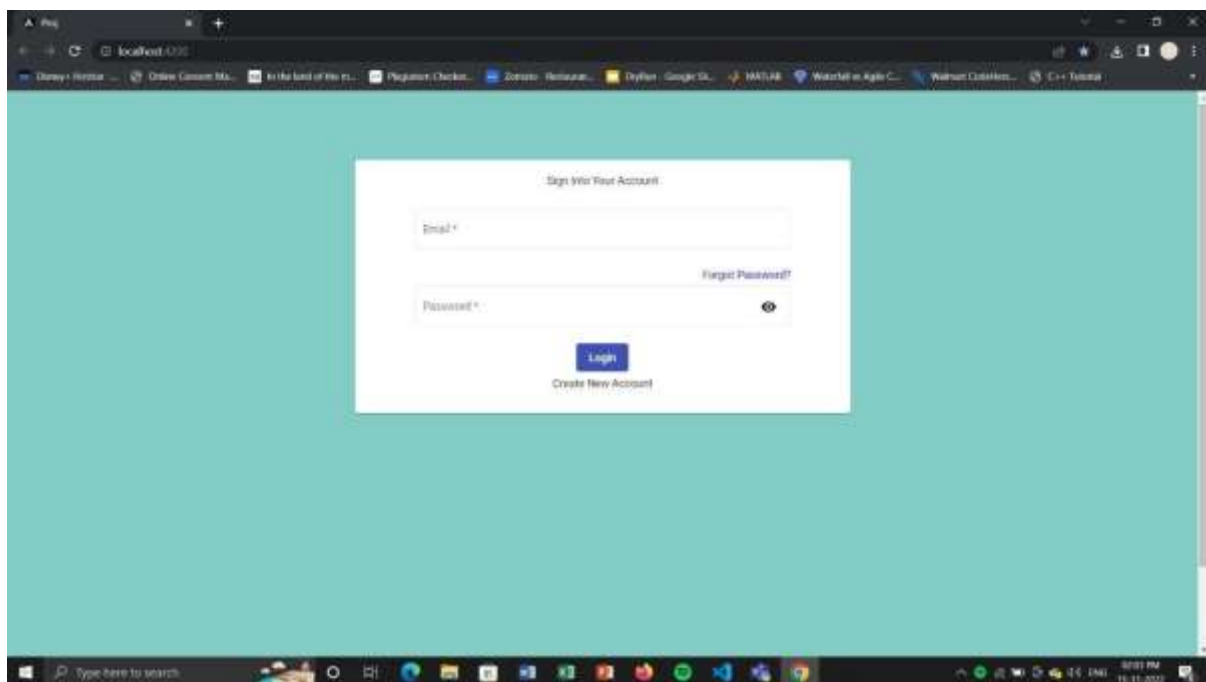


Count of Security



10. RESULTS

10.1 Performance Metrics



11.MERITS AND DEMERITS

Merits

- ✓ Weather aberrations can cause physical damage to crops.
- ✓ Help in cut costs.
- ✓ Product higher crop yield.
- ✓ Prevent over or under watering.

Demerit

- ✓ Sudden change in weather cause crop damage

12. CONCLUSION

Weather aberrations can cause physical damage to crops. With help of this project, we can predict in certain environmental conduction which crop should be taken. From the graph of % of production we can determine sowing and harvesting period of particular crop in given temperature and rainfall. This data will continue to enhance farmer efficiency by further enabling them to monitor each plot of land and determine the precise input needed for their crops.

13. FUTURE SCOPE

1. Predict appropriate crop and maximum yield in the climate change.
2. Create an android app.
3. Collection of data, Analysis of it and modification of the algorithm.
- 4.IOT application in agriculture, automation in production line and man free agriculture which is the future of the world, this is the first step of it.
- 5.Find the percentage yield to happen from the match given percentage in terms of % error.

14. APPENDIX

GitHub Link:

Team ID : PNT2022TMID09610

<https://github.com/IBM-EPBL/IBM-Project-26031-1659980327>

Project Demo Link :

https://drive.google.com/file/d/1rmuJNzUMqtp3pPkWt-1ZUr0KZ8b0Wusp/view?usp=share_link