PROFESSIONAL READINESS FOR INNOVATION, EMPLOYABILITY AND ENTREPRENEURSHIP

ESTIMATE THE CROP YIELD USING DATA ANALYTICS

DOMAIN: DATA ANALYTICS

PROJECT REPORT

Submitted by

TEAM ID

PNT2022TMID13120

TEAM MEMBERS

TL – HARSHINI K - 19I319

TM 1 – GHANASIYAA S – 19I313

TM 2 – MOUNIKA K – 19I330

TM 3 – SUKIRTHI S J – 19I354

NOVEMBER 2022

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Project Overview

In GDP of India, more that 19% share is from Agriculture. Agriculture is first and foremost factor which is important for survival. Machine learning (ML) could be a crucial perspective for acquiring real-world and operative solution for crop yield issue. Considering the present system including manual counting, climate smart pest management and satellite imagery, the result obtained arent really accurate. This paper focuses mainly on predicting the yield of the crop by applying various machine learning techniques.

Human survival depends on agriculture because it provides basic needs. Indian agriculture makes up more than half of the population (55%) and is well known. There are bottlenecks to increasing crop production in India because climatic conditions vary. It has become increasingly difficult to achieve target crop yields in agribusiness. It is necessary to consider a number of factors that directly affect the production and productivity of crops. In agriculture, crop yield prediction is an essential factor.

So, It is important to analyze the Crop Production data of Indian Agriculture market. This project is aimed to create fruitful visualization using Cognos Analytics on cloud for said data. In this project prediction is done to find the insights from Crop Production data of Indian market.

1.2 Purpose

Crop yield can vary with places, season, area,etc. Farmers face a lot of losses due to poor planning and improper planting of crops. In order to meet the food demand of a growing population, we are currently in dire need of another Green Revolution. Globally, there is a decrease in cultivable land and a decrease in cultivable water resources, making it almost impossible to report higher crop yields. Agro-based big data analytics plays a significant role in increasing crop yields by providing the optimal conditions for plants to grow, reducing yield gaps, and reducing crop damage. As a result, in this project we will be creating a dashboard and analyzing some important visualizations to gain more insight into India's crop production.

Based on the prediction, farmers can plan their harvest and sale for the harvest. It helps them to save money and increase their income.

2. LITERATURE SURVEY

2.1 Existing problem

As a result of climate change, extreme weather events have reduced food yields significantly. Crop plants are typically delicate. Due to their high yields rather than their ability to withstand stress. There are four main factors that affect crop production: soil fertility, water availability, climate, and diseases or pests. Without technology, it can be quite challenging to understand or estimate the patterns. The solution must be able to adapt to the changes and offer the anticipated solution in an easy-to-understand manner to end users.

2.2 References

[1] Van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture*, *177*, 105709.

[2] Nigam, A., Garg, S., Agrawal, A., & Agrawal, P. (2019, November). Crop yield prediction using machine learning algorithms. In *2019 Fifth International Conference on Image Information Processing (ICIIP)* (pp. 125-130). IEEE.

[3] Medar, R., Rajpurohit, V. S., & Shweta, S. (2019, March). Crop yield prediction using machine learning techniques. In *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)* (pp. 1-5). IEEE.

[4] Champaneri, M., Chachpara, D., Chandvidkar, C., & Rathod, M. (2016). Crop yield prediction using machine learning. *Technology*, *9*, 38.

[5] Crane-Droesch, A. (2018). Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. *Environmental Research Letters*, *13*(11), 114003.

2.3 Problem Statement Definition

As one of the top countries for crop production, India is a major source of income from crop production. By analyzing some meaningful visualizations and creating a dashboard, we will gain most of the insights into crop production in India. Using the climatic conditions and areas of different states in India, crop yields can be predicted. Cultivating the right crop at the right time enables farmers to yield maximum yields.

For ***"ESTIMATE THE CROP YIELD USING DATA ANALYTICS"*** project, the dataset named "crop_production.csv" is used.The "crop_production.csv" dataset contains 246092 records. This Dataset contains 7 Attributes. The attributes of the given dataset are:-

- State_Name
- District_Name
- Crop_Year
- Season

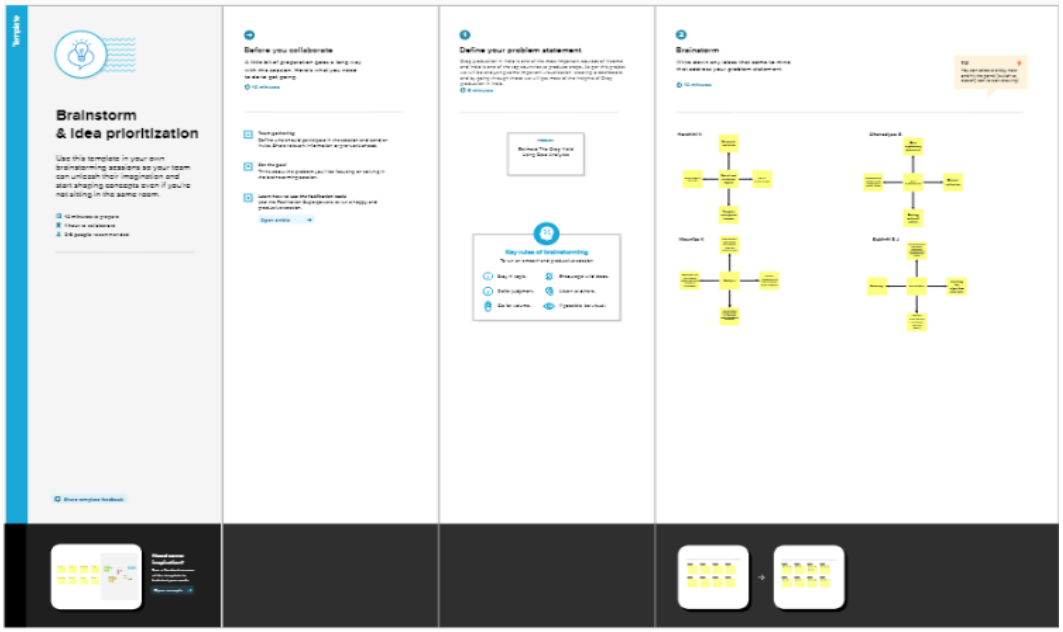
- Crop
- Area
- Production

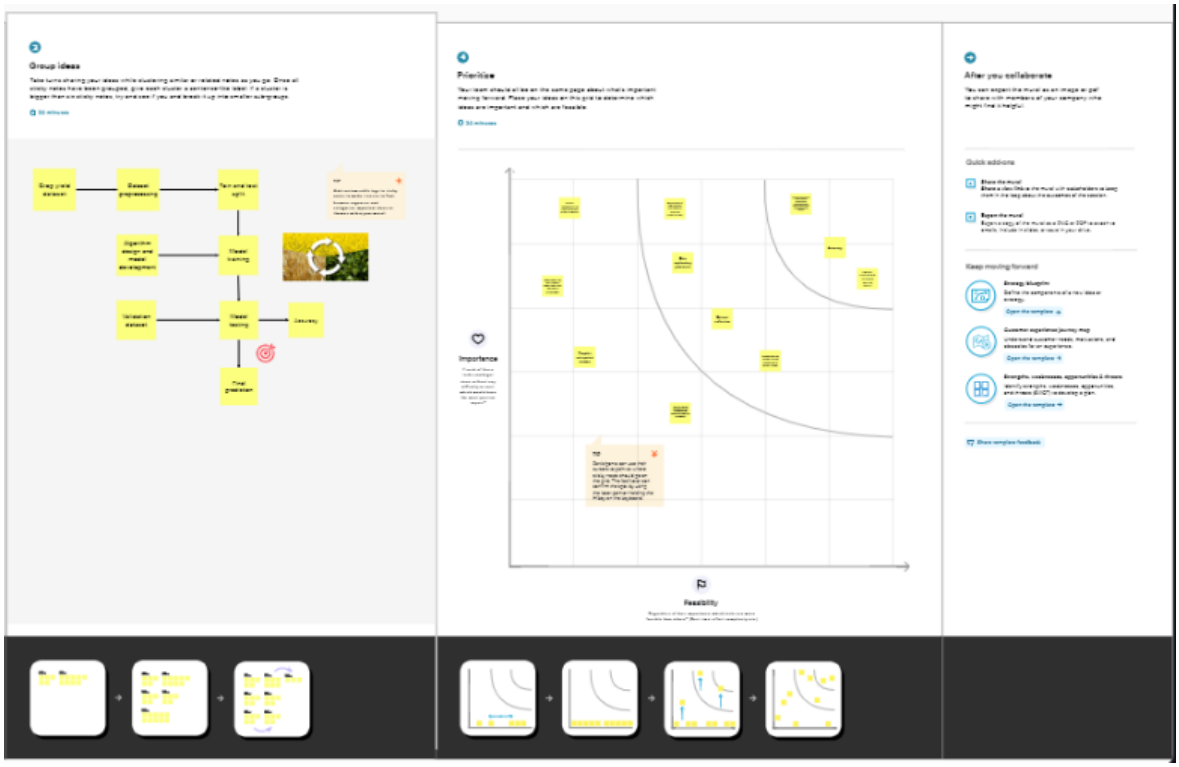
Based on these attributes, crop yield has to be predicted.

3. IDEATION & PROPOSED SOLUTION

3.1Empathy Map Canvas

3.2Ideation & Brainstorming





Before the analysis, the data preprocessing procedures such as missing value analysis, smoothing noisy data and data standardization were applied on the crop dataset to produce reliable data.

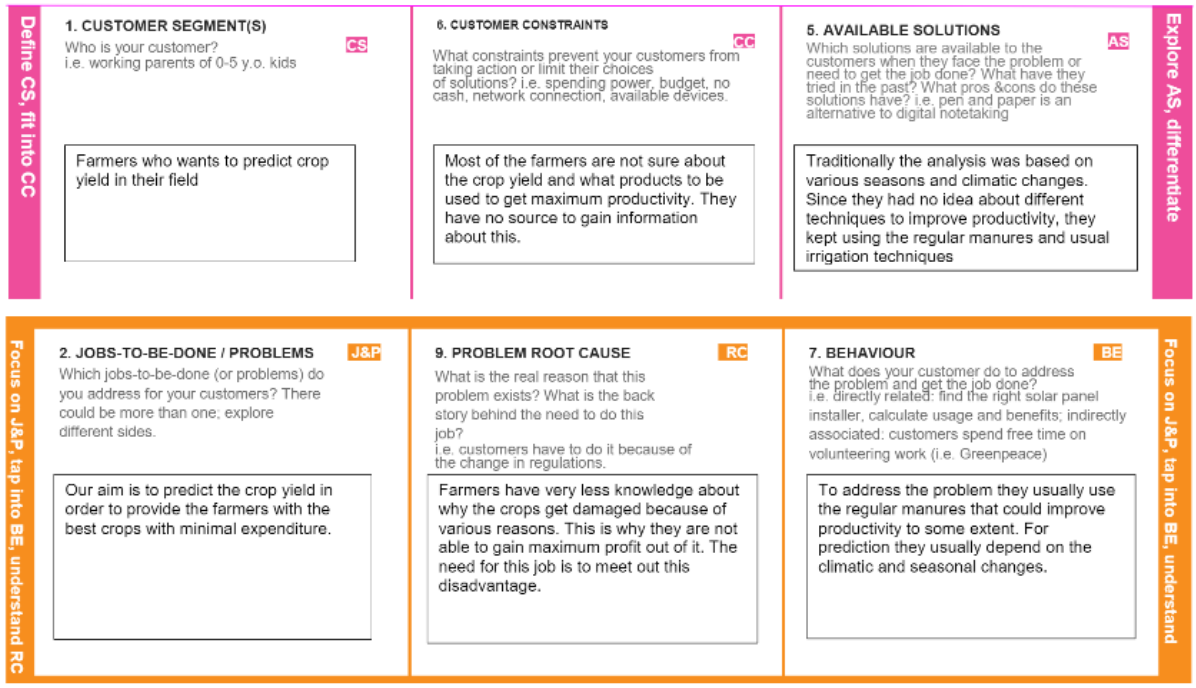
Then the different visualizations are developed using IBM Cognos Watson. The visualization in IBM Cognos Watson automatically analyses crop dataset and dashboard is generated and the dashboard can be customized based on the user requirements. The dashboard is displayed on an Application UI.

3.3 Proposed Solution

S.NO	Parameter	Description
1.	Problem Statement (Problemto be solved)	The ultimate aim of this project is to predict the crop yield with the help of certain data analysis techniques. This can be extremely useful to the farmers so that they can plan their harvests accordingly.
2.	Idea / Solution description	Based on the analysis made on the crop yields, it is feasible to create a mobile or web application which contain details not only about the harvest periods but also about the nature of crops, properties of soil and so on.

3.	Novelty / Uniqueness	This analysis can bring out the goodness of every type of soil and helps us understand the crops better.
4.	Social Impact / CustomerSatisfaction	By getting to know about the crop and what time is suitable for harvest, we can make sure that we never hit famine and protect people from running out of food.
5.	Business Model (RevenueModel)	This analysis helps farmers to understand their crops and based on that they can decide whether or not to invest on a particular crop variety. This helps in improving the crop production and thereby increases the revenue of the farmers.
6.	Scalability of the Solution	Once the basic prediction model is designed it can be used to perform analysis on any type of soil and on any crop variety.

3.4 Problem Solution Fit



4. REQUIREMENT ANALYSIS

4.1 Functional requirements:

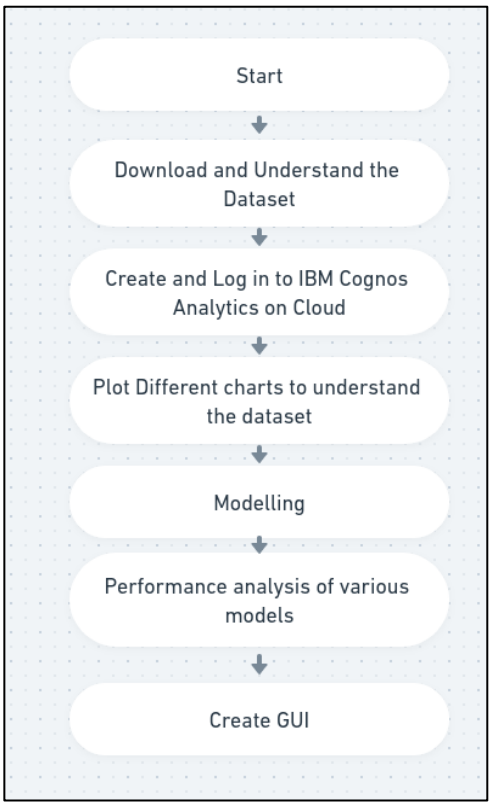
Functional requirement	Description
Registration	Registers a new user through registration form
Google authentication	New user can get added even by using a google account
Login	Lets the registered user to login to the portal
Take in the required data	This takes in the required data from the user
Estimation	A prediction of crop yield is done based on the current data
Analysis	An analysis is done on the given data to gain useful insights

4.2 Non-functional requirements:

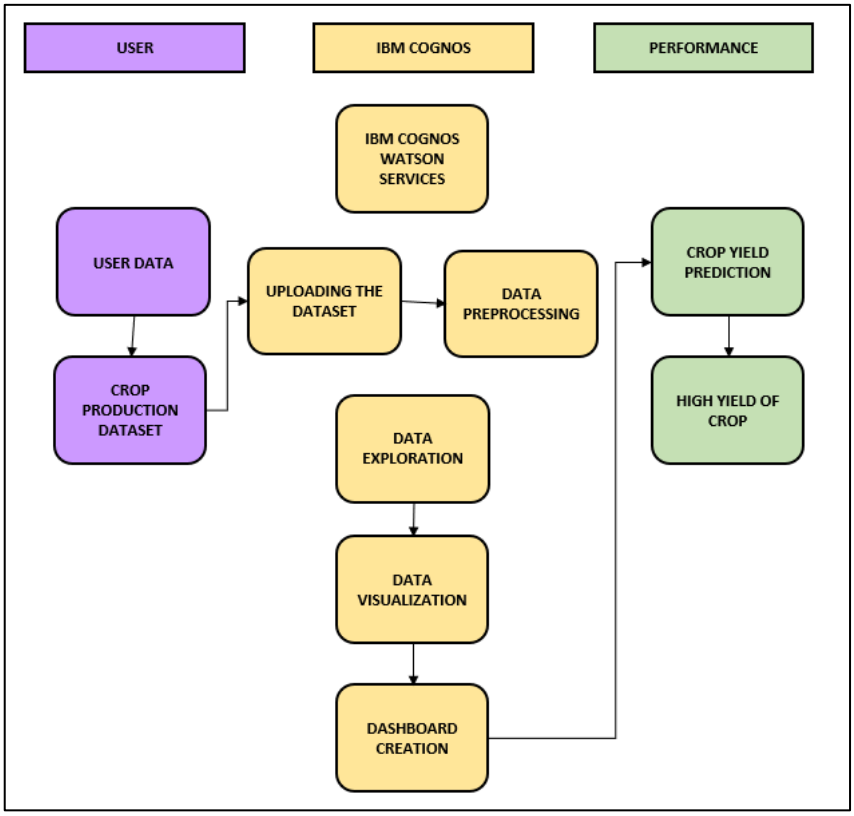
Non-functional requirement	Description
Performance	The software should provide us good performance
Reliability	The UI should be user friendly and easily understandable
Availability	It should be available for access at any time from anywhere
Scalability	The software should be flexible and other developers must be able to improve its capabilities

5. PROJECT DESIGN

5.1 Data Flow Diagrams



5.2Solution & Technical Architecture



5.3 User Stories

Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Registration	USN-1	As a user, I can register for by entering my Agri - id card and request.	2	High	Harshini, Ghanasiyaa
	USN-3	As a user, I can register for the application through Gmail	2	Medium	Sukirthi Mounika
Login	USN-4	As a user, I can Call and request or Approach for dataset	4	High	Harshini, Ghanasiyaa
Working with the Dataset	USN-5	To work on the given dataset, Understand the Dataset.	2	High	Sukirthi Mounika
	USN-6	Load the dataset to Cloud platform then Build the requiredVisualizations.	10	High	Harshini, Ghanasiyaa
Data Visualization Chart	USN-7	Using the Crop production in Indian dataset, create various graphsand charts to highlight the insights and visualizations. *Build a Visualization to showcase Average Crop Production bySeasons.	4	Medium	Sukirthi Mounika
		*Showcase the Yearly usage of Area in Crop Production.	4	Medium	Harshini, Mounika

6. PROJECT PLANNING & SCHEDULING

6.1Sprint Planning & Estimation

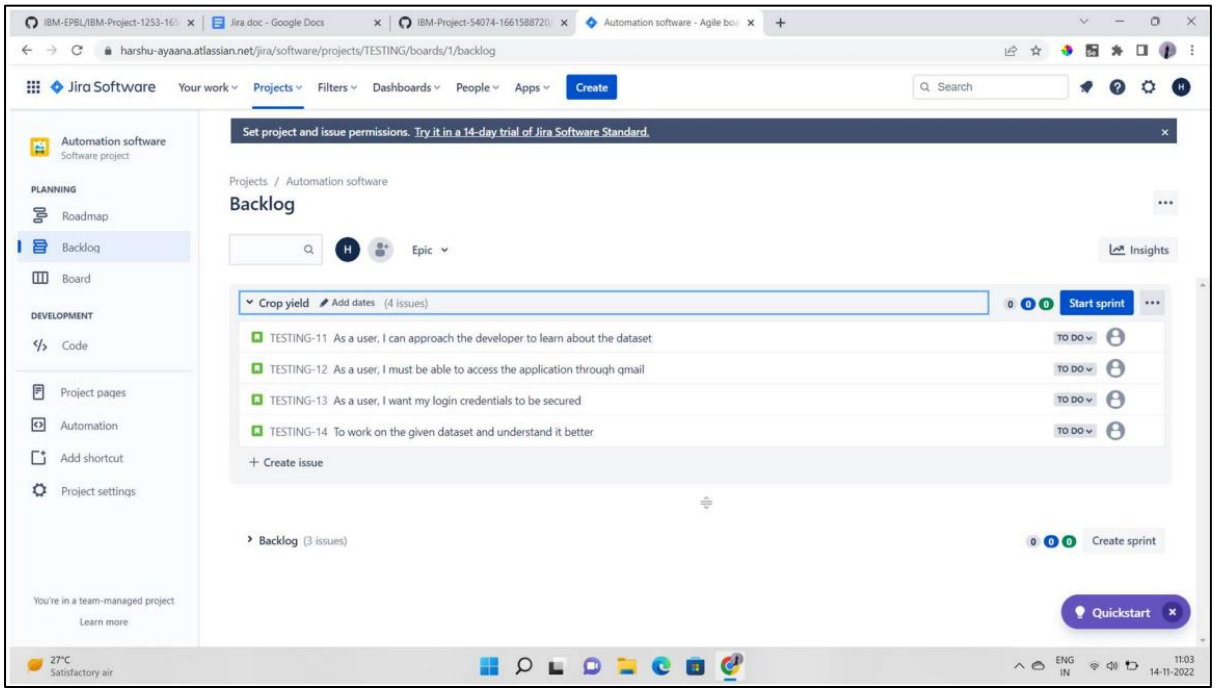
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Working with dataset	USN-1	To work on the given dataset, Understand the Dataset.	2	High	Ghanasiyaa S, Sukirthi S J
Sprint-1		USN-2	Load the dataset to Cloud platform then Build the required Visualizations.	3	High	Ghanasiyaa S, Sukirthi S J
Sprint-2	Data Visualization Chart	USN-3	Using the Crop production in Indian dataset, create various graphs and charts to highlight the insights and visualizations. *Build a Visualization to showcase Average Crop Production by Seasons.	4	Medium	Ghanasiyaa S, Sukirthi S J, Harshini K
		USN-4	*Build a visualization to show case top 10 States in Crop Yield Production by Area.	4	Medium	Mounika K, Sukirthi S J
		USN-5	*Build the required Visualization to showcase the Crop Production by State.	4	Medium	Harshini K, Mounika K
		USN-6	*Build Visual analytics to represent the Sates with Seasonal Crop Production using a Text representation.	4	Medium	Harshini K, Sukirthi S J
Sprint-3	Creating the dashboard	USN-7	Create the Dashboard by using the created visualizations.	20	High	Ghanasiyaa S, Mounika K, Sukirthi S J, Harshini K
Sprint-4	Export the Analytics	USN-8	Export the created Dashboard	20	High	Mounika K, Harshini K

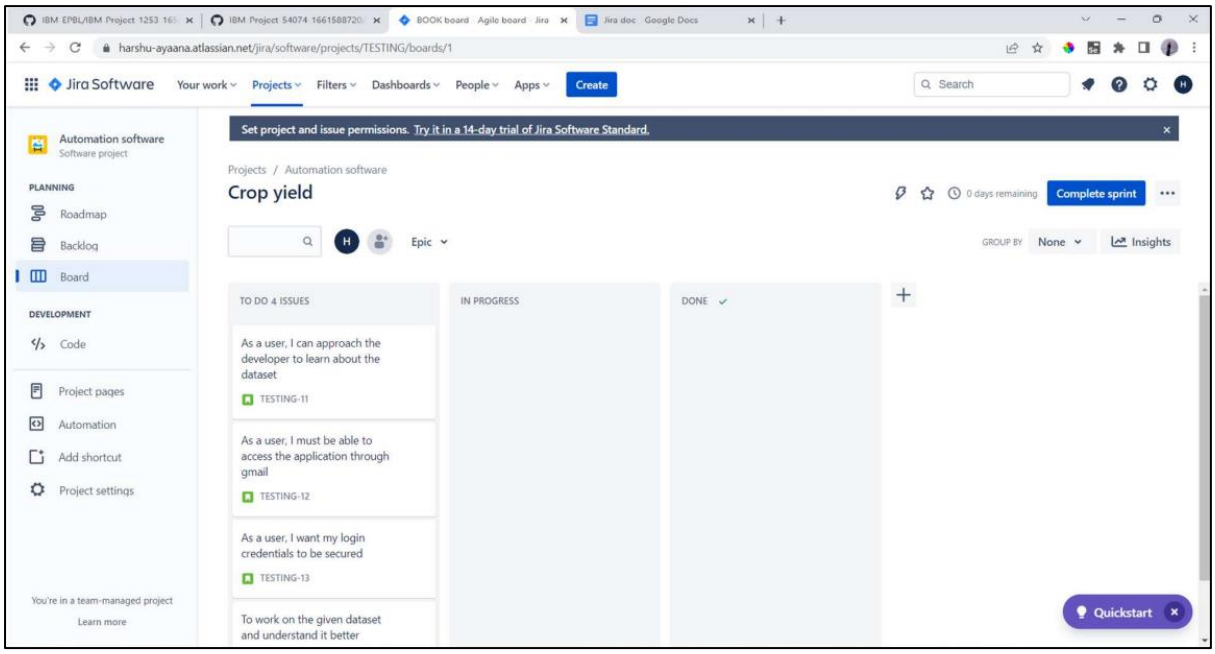
6.2Sprint Delivery Schedule

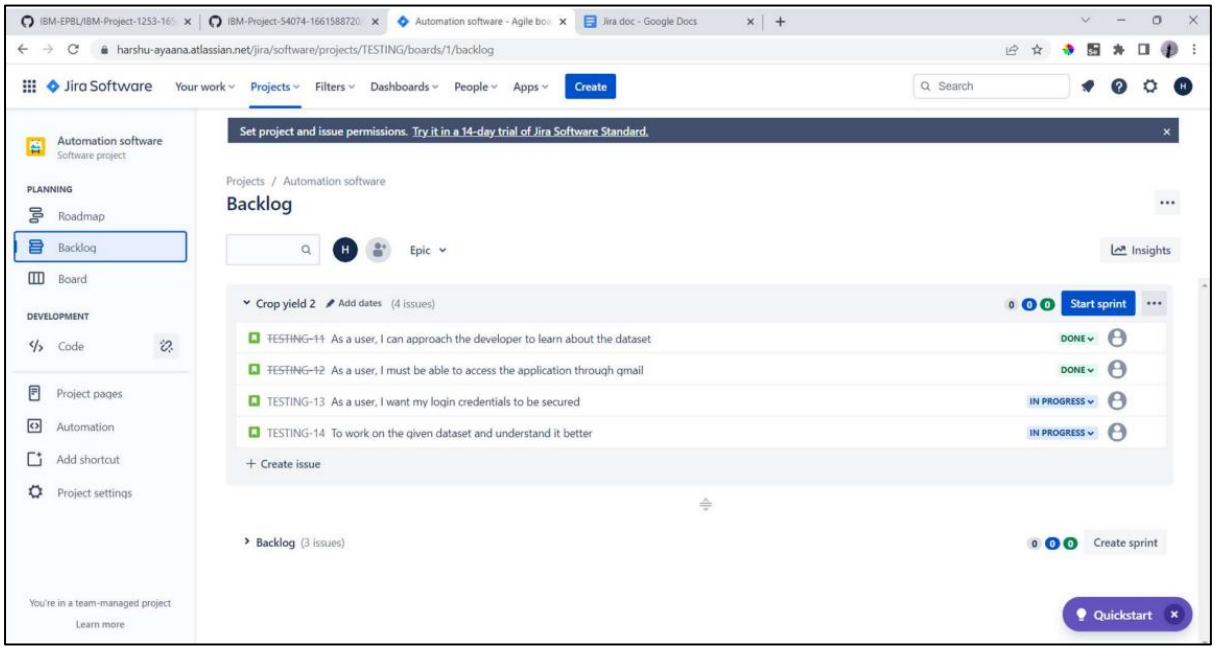
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022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

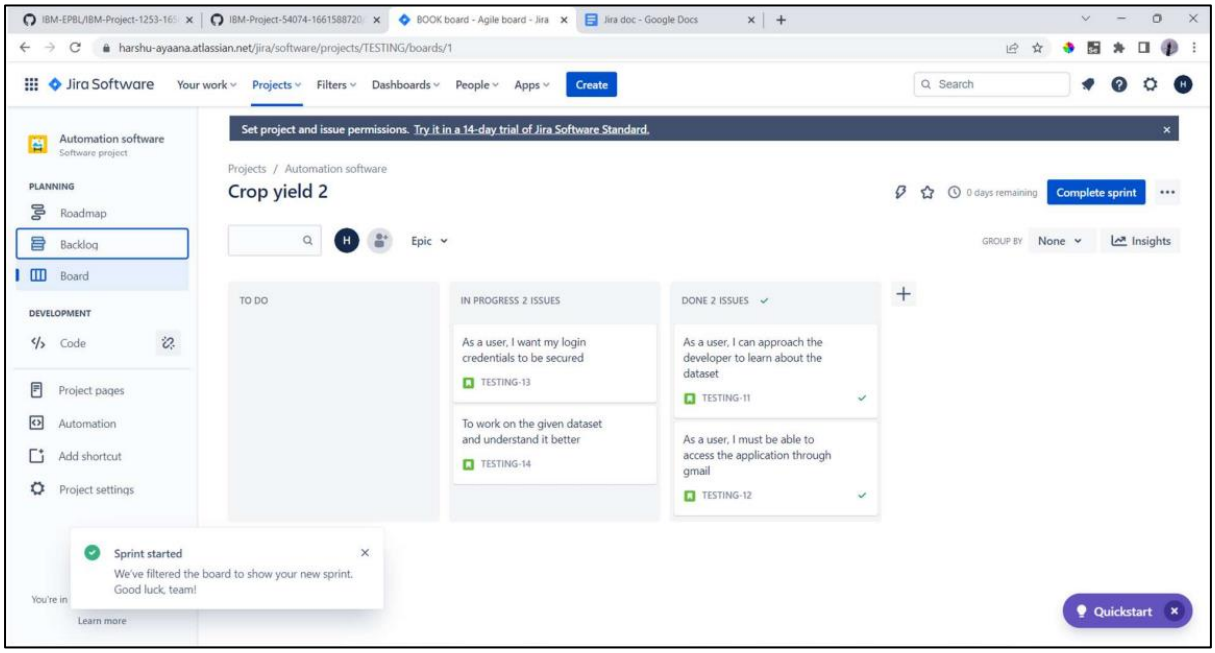
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 ov 2022

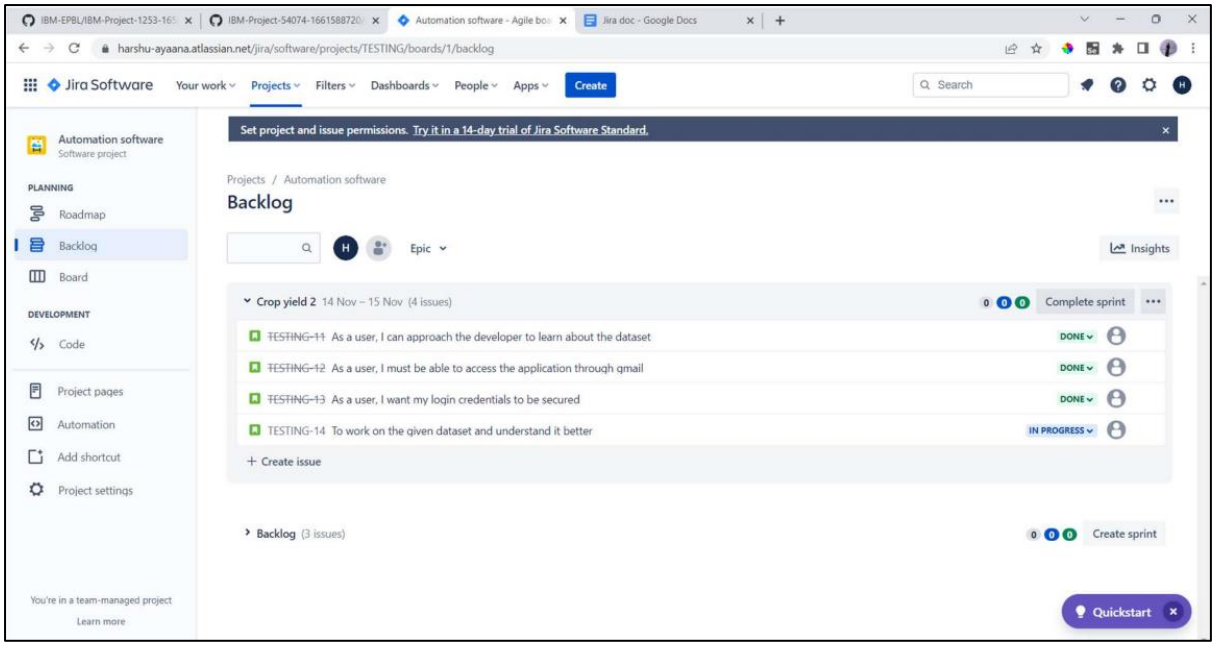
6.3Reports from JIRA

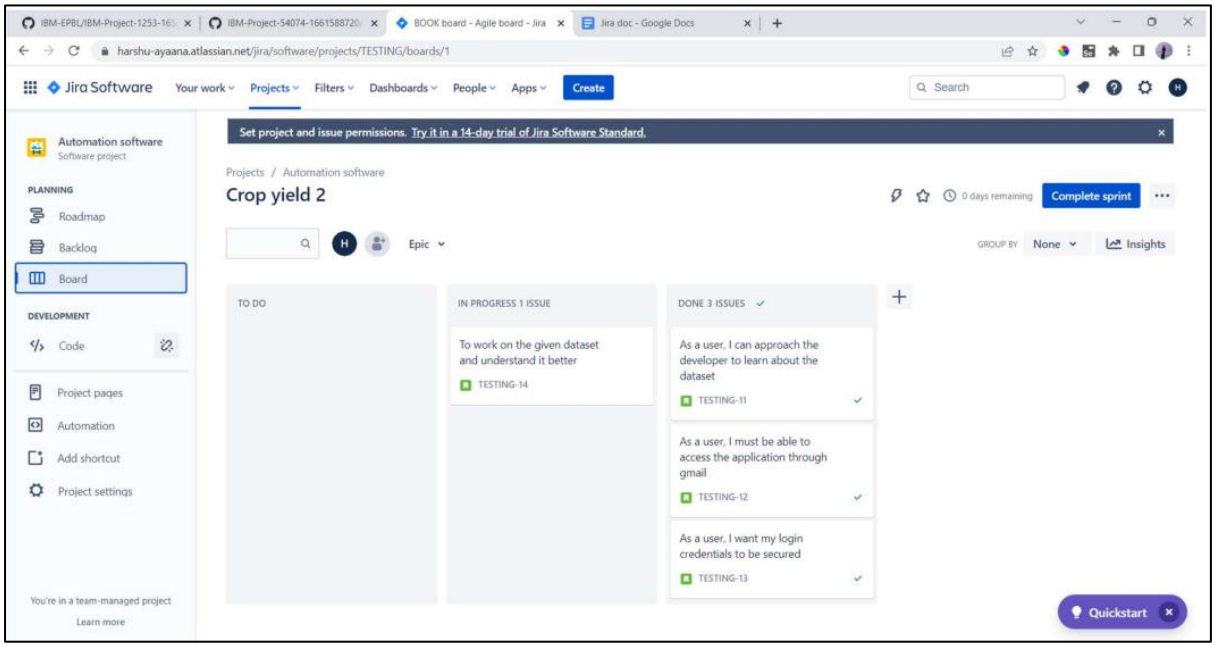


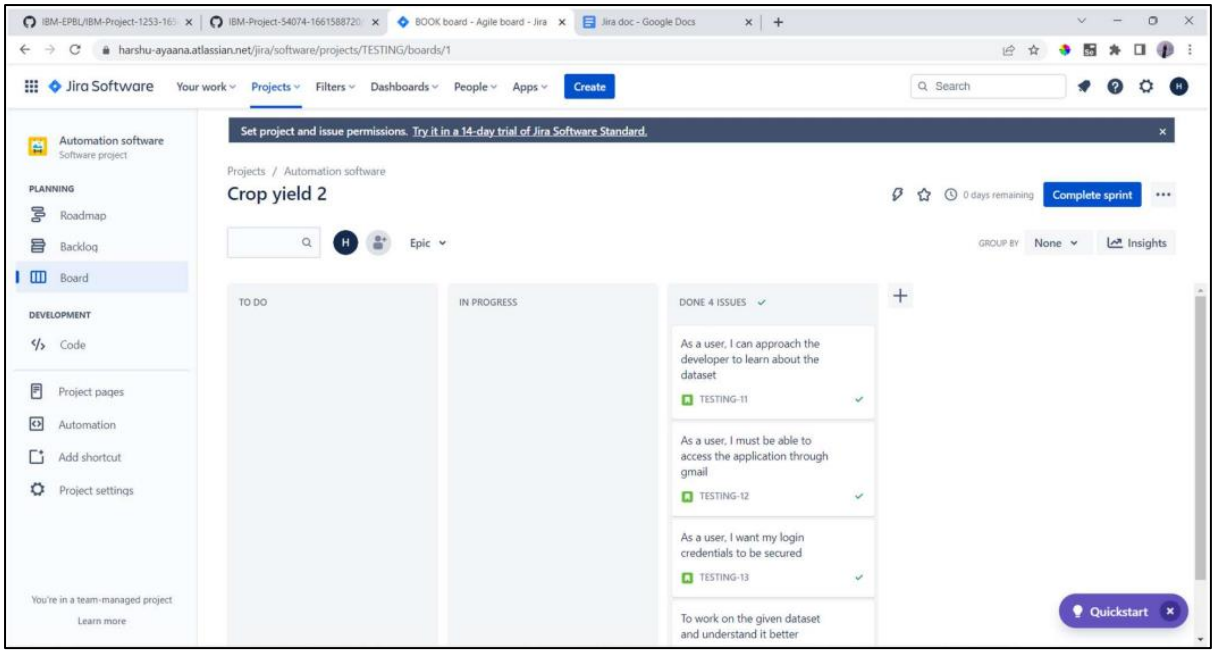


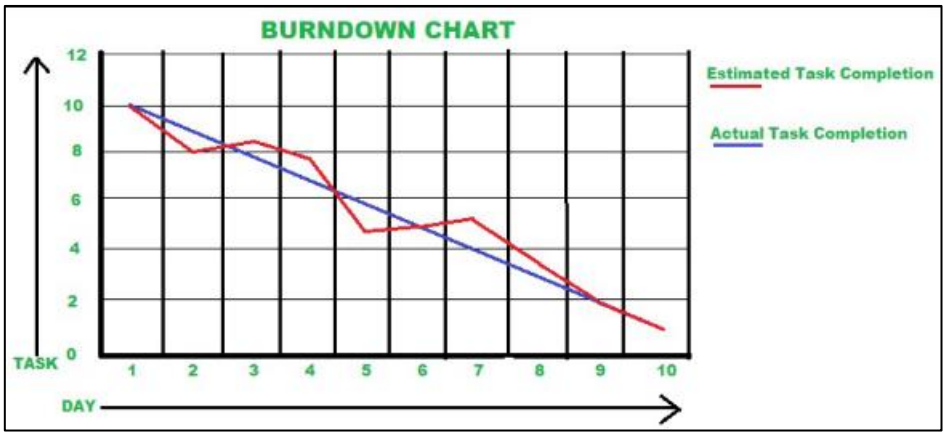

Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Automation software
Software project
PLANNING
Roadmap
Backlog
Board
DEVELOPMENT
Code
Project pages
Automation
Add shortcut
Project settings
You're in a team-managed project
Learn more
Set project and issue permissions. Try it in a 14-day trial of Jira Software Standard.
Projects / Automation software
Backlog
Epic
Insights
Crop yield 2
Add dates
(4 issues)
Start sprint
TESTING-11 As a user, I can approach the developer to learn about the dataset
DONE
TESTING-12 As a user, I must be able to access the application through gmail
DONE
TESTING-13 As a user, I want my login credentials to be secured
IN PROGRESS
TESTING-14 To work on the given dataset and understand it better
IN PROGRESS
Create issue
Backlog (3 issues)
Create sprint
Quickstart


Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Automation software
Software project
PLANNING
Roadmap
Backlog
Board
DEVELOPMENT
Code
Project pages
Automation
Add shortcut
Project settings
Set project and issue permissions. Try it in a 14-day trial of Jira Software Standard.
Projects / Automation software
Crop yield 2
0 days remaining
Complete sprint
Epic
GROUP BY
None
Insights
TO DO
IN PROGRESS 2 ISSUES
As a user, I want my login credentials to be secured
TESTING-13
To work on the given dataset and understand it better
TESTING-14
DONE 2 ISSUES
As a user, I can approach the developer to learn about the dataset
TESTING-11
As a user, I must be able to access the application through gmail
TESTING-12
Sprint started
We've filtered the board to show your new sprint. Good luck, team!
Learn more
Quickstart


harshu-ayaana.atlassian.net/jira/software/projects/TESTING/boards/1/backlog
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Automation software
Software project
PLANNING
Roadmap
Backlog
Board
DEVELOPMENT
Code
Project pages
Automation
Add shortcut
Project settings
You're in a team-managed project
Learn more
Set project and issue permissions. Try it in a 14-day trial of Jira Software Standard.
Projects / Automation software
Backlog
Epic
Insights
Crop yield 2 14 Nov – 15 Nov (4 issues)
Complete sprint
TESTING-11 As a user, I can approach the developer to learn about the dataset
DONE
TESTING-12 As a user, I must be able to access the application through gmail
DONE
TESTING-13 As a user, I want my login credentials to be secured
DONE
TESTING-14 To work on the given dataset and understand it better
IN PROGRESS
Create issue
Backlog (3 issues)
Create sprint
Quickstart


harshu-ayaana.atlassian.net/jira/software/projects/TESTING/boards/1
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Automation software
Software project
PLANNING
Roadmap
Backlog
Board
DEVELOPMENT
Code
Project pages
Automation
Add shortcut
Project settings
You're in a team-managed project
Learn more
Set project and issue permissions. Try it in a 14-day trial of Jira Software Standard.
Projects / Automation software
Crop yield 2
0 days remaining
Complete sprint
Epic
GROUP BY
None
Insights
TO DO
IN PROGRESS 1 ISSUE
To work on the given dataset and understand it better
TESTING-14
DONE 3 ISSUES
As a user, I can approach the developer to learn about the dataset
TESTING-11
As a user, I must be able to access the application through gmail
TESTING-12
As a user, I want my login credentials to be secured
TESTING-13
Quickstart


Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Automation software
Software project
PLANNING
Roadmap
Backlog
Board
DEVELOPMENT
Code
Project pages
Automation
Add shortcut
Project settings
You're in a team-managed project
Learn more
Set project and issue permissions. Try it in a 14-day trial of Jira Software Standard.
Projects / Automation software
Crop yield 2
0 days remaining
Complete sprint
Epic
GROUP BY
None
Insights
TO DO
IN PROGRESS
DONE 4 ISSUES
As a user, I can approach the developer to learn about the dataset
TESTING-11
As a user, I must be able to access the application through gmail
TESTING-12
As a user, I want my login credentials to be secured
TESTING-13
To work on the given dataset and understand it better
Quickstart


BURNDOWN CHART
12
10
8
6
4
2
0
TASK
1
2
3
4
5
6
7
8
9
10
DAY
Estimated Task Completion
Actual Task Completion

7. CODING & SOLUTIONING

7.1 Feature 1

The model used for the project is Random Forest. Random forests or random decision forests is an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, the mean or average prediction of the individual trees is returned. Here Random Forest Regressor is used. Following is the code used in the project.

```
from sklearn.ensemble import RandomForestRegressor
model = RandomForestRegressor()
model.fit(x_train,y_train)
preds = model.predict(x_test)
```

7.2 Feature 2

Another model called linear Regression is used for the project. Linear Regression is a machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on the kind of relationship between dependent and independent variables they are considering, and the number of independent variables getting used. Following is the code used in the project.

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

7.3 coding

```
<!Doctype HTML>
<html>
<head>
<title></title>
```

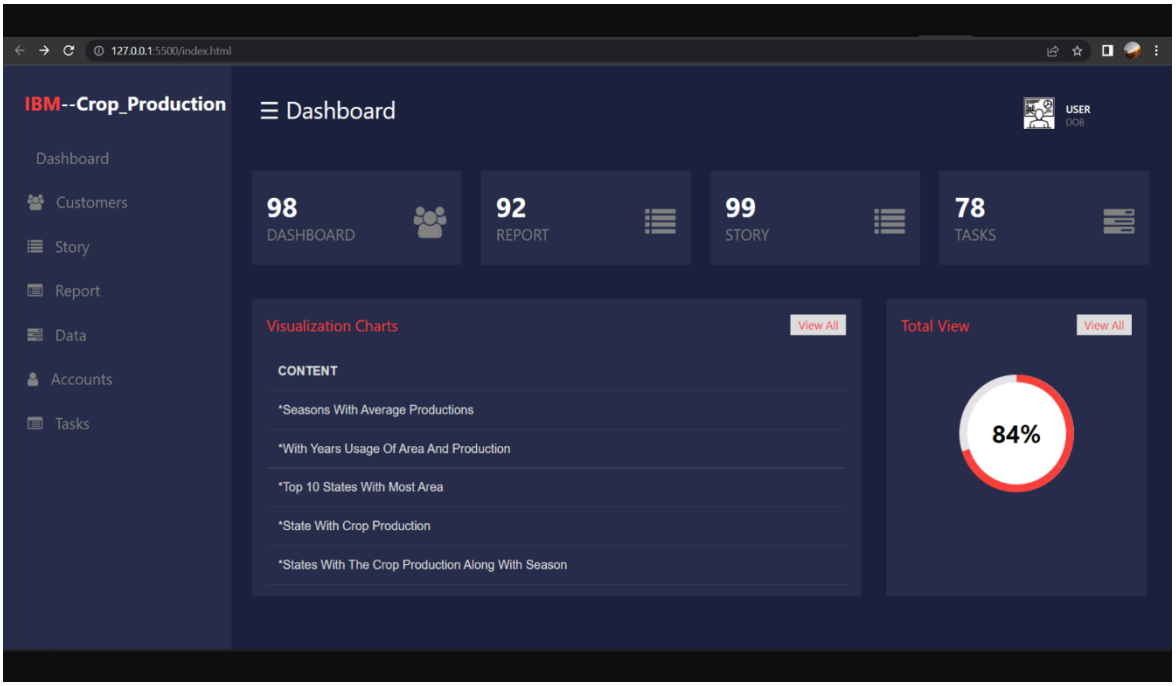
```
<link rel="stylesheet" href="style.css" type="text/css"/>
<link rel="stylesheet" href="https://cdnjs.cloudflare.com/ajax/libs/font-awesome/4.7.0/css/font-awesome.min.css">
</head>
<body>
<div id="mySidenav" class="sidenav">
<p class="logo"><span>IBM</span>--Crop_Production</p>
<ahref="https://drive.google.com/file/d/1dRhLaS7ZeC1rDN78Ukhqn5R1Q_SxUWX/view?usp=sharing"></i>   Dashboard</a>
<a href="#"class="icon-a"><i class="fa fa-users icons"></i>   Customers</a>
<a href="https://drive.google.com/file/d/1_pO8Y2guW-m_xCPBROiZNyfUasdjffC5/view?usp=sharing"class="icon-a"><i class="fa fa-list icons"></i>   Story</a>
<ahref="https://drive.google.com/file/d/1yq32laQkjzzDdOVbl0UeFOig10rdTpU4/view?usp=sharing"class="icon-a"><i class="fa fa-list-alt icons"></i>   Report</a>
<a href="https://www.kaggle.com/datasets/abhinand05/crop-production-in-india"class="icon-a"><i class="fa fa-tasks icons"></i>   Data</a>
<a href="#"class="icon-a"><i class="fa fa-user icons"></i>   Accounts</a>
<a href="#"class="icon-a"><i class="fa fa-list-alt icons"></i>   Tasks</a>
</div>
<div id="main">
<div class="head">
<div class="col-div-6">
<span style="font-size:30px;cursor:pointer; color: white;" class="nav" >&#9776; Dashboard</span>
<span style="font-size:30px;cursor:pointer; color: white;" class="nav2" >&#9776; Dashboard</span>
</div>
<div class="col-div-6">
<div class="profile">
<img src="images/user.png" class="pro-img" />
<p>USER <span>DOB</span></p>
</div>
</div>
```

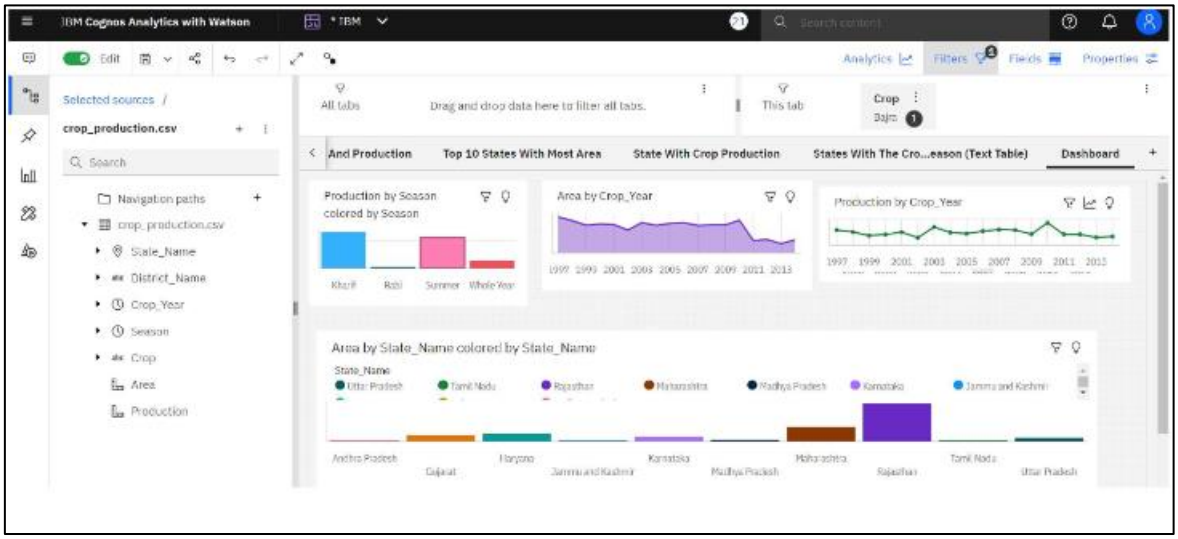
```
<div class="clearfix"></div>
</div>
<div class="clearfix"></div>
<br/>
<div class="col-div-3">
<div class="box">
<p>98<br/><span>DASHBOARD</span></p>
<i class="fa fa-users box-icon"></i>
</div>
</div>
<div class="col-div-3">
<div class="box">
<p>92<br/><span>REPORT</span></p>
<i class="fa fa-list box-icon"></i>
</div>
</div>
<div class="col-div-3">
<div class="box">
<p>99<br/><span>STORY</span></p>
<i class="fa fa-list box-icon"></i>
</div>
</div>
<div class="col-div-3">
<div class="box">
<p>78<br/><span>TASKS</span></p>
<i class="fa fa-tasks box-icon"></i>
</div>
</div>
<div class="clearfix"></div>
<br/><br/>
<div class="col-div-8">
<div class="box-8">
<div class="content-box">
<p>Visualization Charts <span>View All</span></p>
```

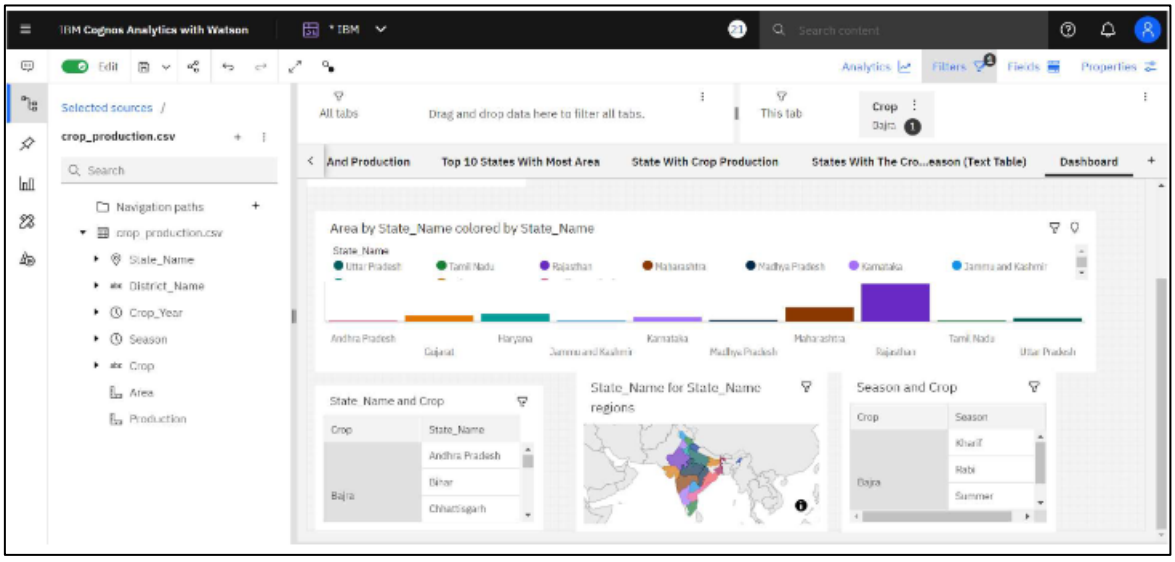
```
<br/>
<table>
<tr>
<th>CONTENT</th>
</tr>
<tr>
<td>*Seasons With Average Productions</td>
</tr>
<tr>
<td>*With Years Usage Of Area And Production</td>
</tr>
<tr>
<td>*Top 10 States With Most Area</td>
</tr>
<tr>
<td>*State With Crop Production</td>
</tr>
<tr>
<td>*States With The Crop Production Along With Season </td>
</tr>
</table>
</div>
</div>
</div>
<div class="col-div-4">
<div class="box-4">
<div class="content-box">
<p>Total View <span>View All</span></p>
<div class="circle-wrap">
<div class="circle">
<div class="mask full">
<div class="fill"></div>
</div>
<div class="mask half">
```

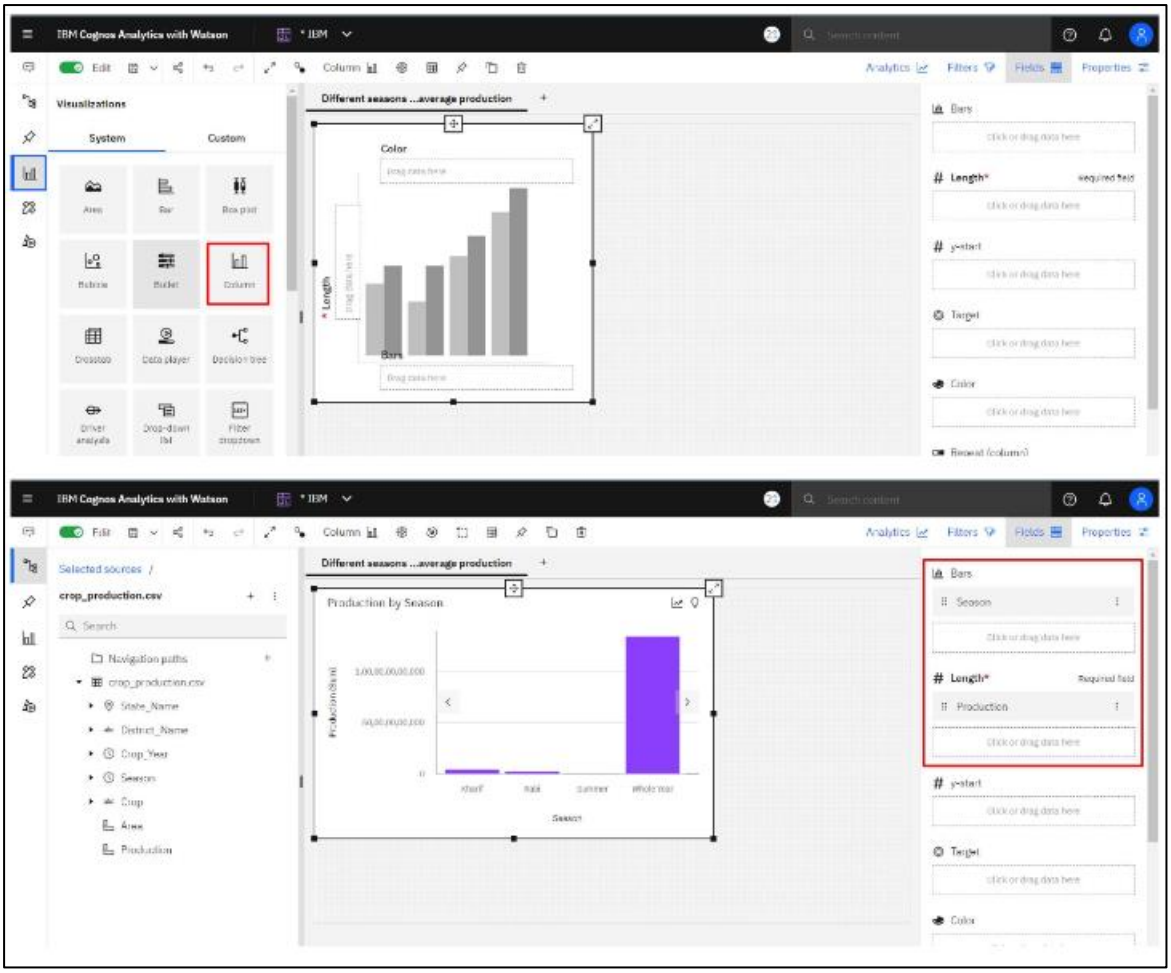
```
<div class="fill"></div>
</div>
<div class="inside-circle"> 84% </div>
</div>
</div>
</div>
<div>
</div>
<div class="clearfix"></div>
</div>
<script src="https://ajax.googleapis.com/ajax/libs/jquery/3.5.1/jquery.min.js"></script>
<script>
 $(".nav").click(function(){
    $("#mySidenav").css('width','70px');
    $("#main").css('margin-left','70px');
    $(".logo").css('visibility', 'hidden');
    $(".logo span").css('visibility', 'visible');
     $(".logo span").css('margin-left', '-10px');
     $(".icon-a").css('visibility', 'hidden');
     $(".icons").css('visibility', 'visible');
     $(".icons").css('margin-left', '-8px');
      $(".nav").css('display','none');
      $(".nav2").css('display','block');
  });
$(".nav2").click(function(){
    $("#mySidenav").css('width','300px');
    $("#main").css('margin-left','300px');
    $(".logo").css('visibility', 'visible');
     $(".icon-a").css('visibility', 'visible');
     $(".icons").css('visibility', 'visible');
     $(".nav").css('display','block');
      $(".nav2").css('display','none');
 });
</script></body></html>
```

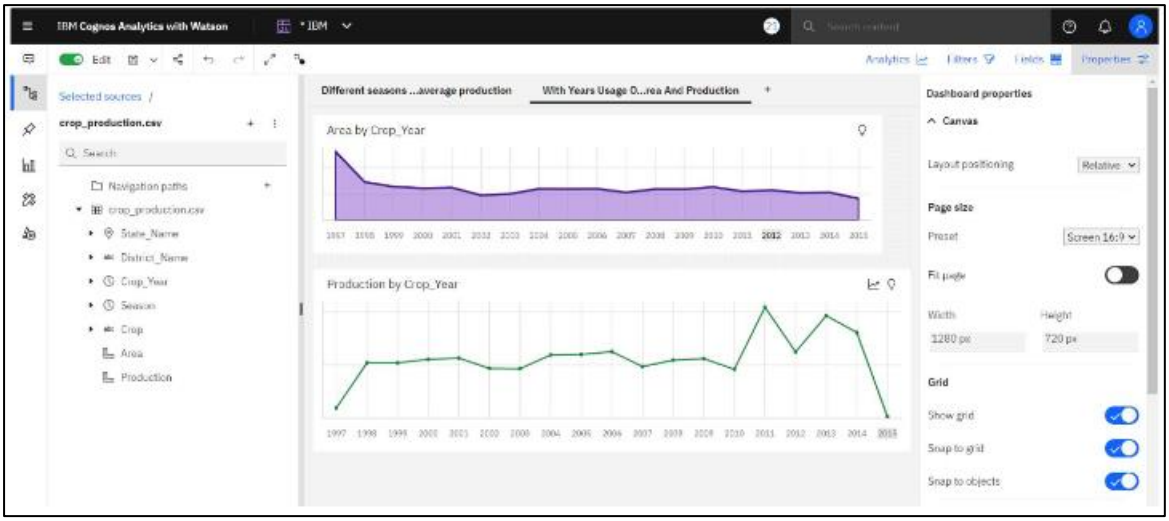
7.4 Report Creation

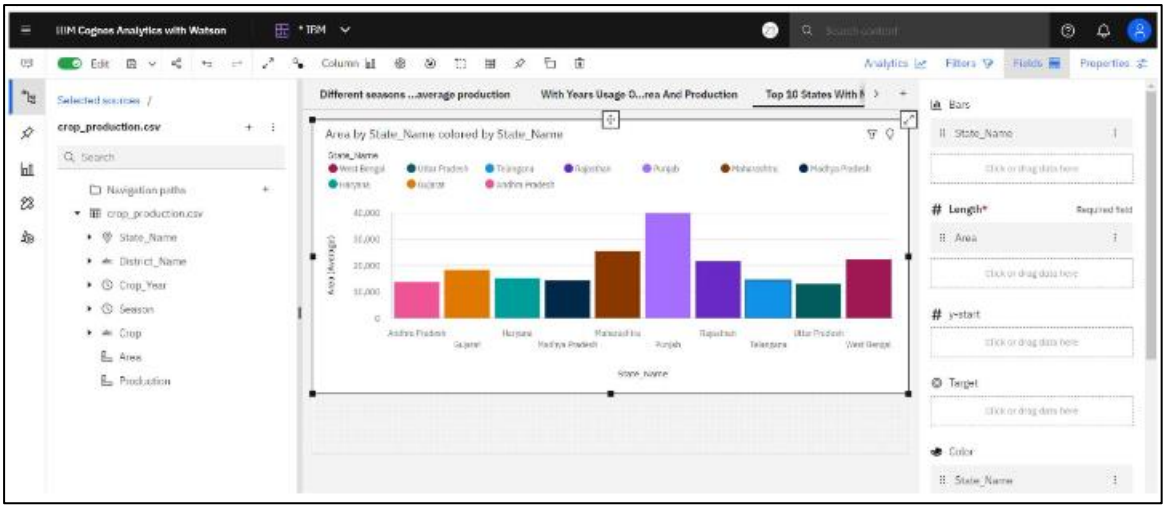


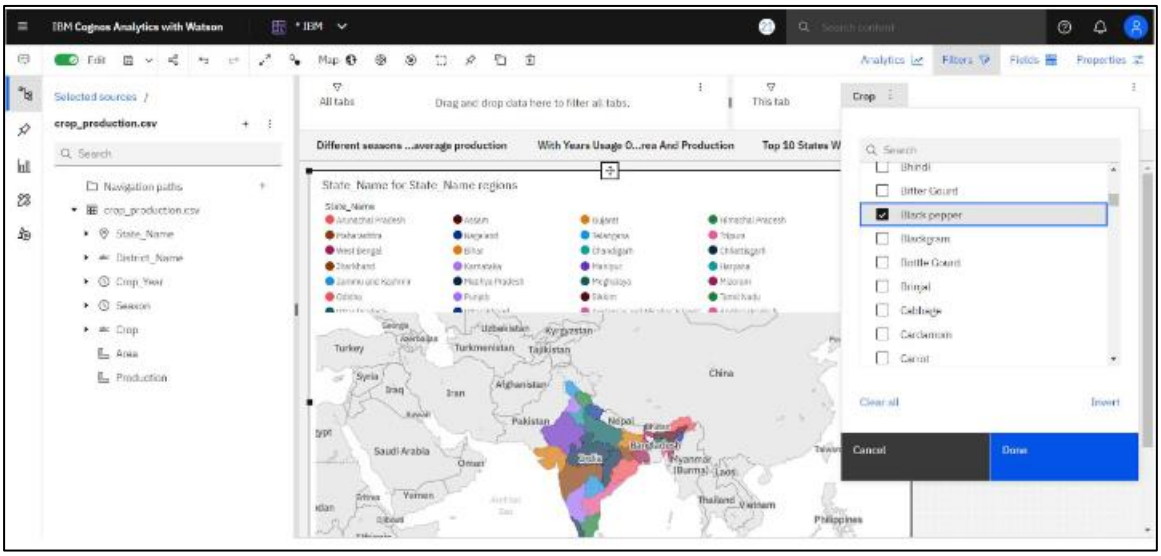


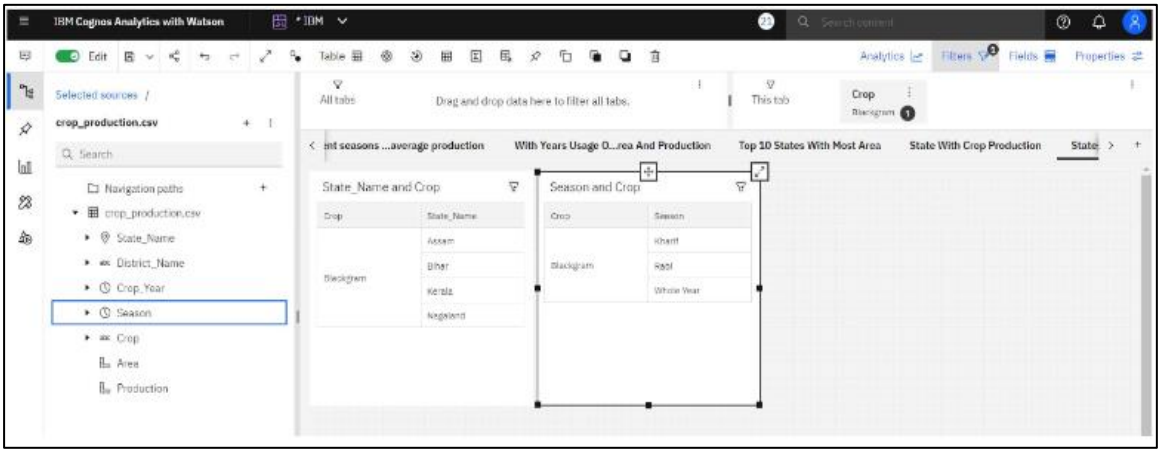



IBM Cognos Analytics with Watson
Visualizations
System
Custom
Different seasons ...average production
Color
Length
Bars
Production by Season
Selected sources
crop_production.csv
Navigation paths
State_Name
District_Name
Crop_Year
Season
Crop
Area
Production
Bars
Length*
Target
Color


IBM Cognos Analytics with Watson
Different seasons ...average production
With Years Usage O...rea And Production
Area by Crop_Year
Production by Crop_Year
Dashboard properties
Canvas
Layout positioning
Relative
Page size
Preset
Screen 16:9
Width
Height
1280 px
720 px
Grid
Show grid
Snap to grid
Snap to objects


IBM Cognos Analytics with Watson
Edit
Column
Analytics
Filters
Fields
Properties
Selected sources
crop_production.csv
Search
Navigation paths
crop_production.csv
State_Name
District_Name
Crop_Year
Season
Crop
Area
Production
Different seasons ...average production
With Years Usage O...rea And Production
Top 10 States With
Area by State_Name colored by State_Name
Bars
State_Name
Length*
Area
y-start
Target
Color
State_Name


IBM Cognos Analytics with Watson
Edit
Map
Analytics
Filters
Fields
Properties
Selected sources
crop_production.csv
Search
Navigation paths
crop_production.csv
State_Name
District_Name
Crop_Year
Season
Crop
Area
Production
All tabs
Drag and drop data here to filter all tabs.
This tab
Crop
Different seasons ...average production
With Years Usage O...rea And Production
Top 10 States W
State_Name for State_Name regions
Search
Bhindi
Bitter Gourd
Black pepper
Blackgram
Bottle Gourd
Brinjal
Cabbage
Cardamom
Carrot
Clear all
Invert
Cancel
Done


IBM Cognos Analytics with Watson
Edit
Table
Analytics
Filters
Fields
Properties
Selected sources
crop_production.csv
Search
Navigation paths
crop_production.csv
State_Name
District_Name
Crop_Year
Season
Crop
Area
Production
All tabs
Drag and drop data here to filter all tabs.
This tab
Crop
Blackgram
With Years Usage O...rea And Production
Top 10 States With Most Area
State With Crop Production
State_Name and Crop
Crop
State_Name
Blackgram
Assam
Bihar
Kerala
Nagaland
Season and Crop
Crop
Season
Blackgram
Kharif
Rabi
Whole Year

8. TESTING

8.1 Test Cases

The dataset consists of columns like state name, district name, crop year, season, crop and area. By using these features production is determined.

Test cases include various state names, district names, crop year, season, kind of crop and amount of area.

Following is the example of the test cases which can be given to the model.

State Name - Andaman and Nicobar Islands
District Name - SOUTH ANDAMANS
Year - 2010
Season - Kharif
Crop - Rice
area - 3118.00

8.2 User Acceptance Testing

Based on the user's state, district, kind of crop the user wants to estimate, season, area of the user's land the model's output will be given. The user can provide different crop names for the prediction. The user can check the amount of production for various crops and determine which crop to use in the land.

The GUI developed is user friendly and the user can easily enter the values and get the prediction.

9. RESULTS

9.1 Performance Metrics

The performance metrics used is R squared. The coefficient of determination, denoted R2 or r2 and pronounced "R squared", is the proportion of the variation in the dependent variable that is predictable from the independent variable(s).

The following are performance of the different models.

Random Forest Regressor:

```
from sklearn.metrics import r2_score
r = r2_score(y_test,preds)
print("R2score when we predict using Randomn forest is ",r)
```

```
R2score when we predict using Randomn forest is  0.9978428774834669
```

The score is 0.9978

Linear Regression:

```
from sklearn.metrics import mean_squared_error, r2_score
mean_squared_error(y_test,preds)
r2_score(y_test,preds)
```

```
1.0
```

The score is 1.0

10. ADVANTAGES & DISADVANTAGES

Following is the Advantages of Proposed solution.

- Use of varied data sourceslike csv, .txt, json, etc.
- An intuitive and straightforward user-friendly interface.
- Personalized experience.
- Smart search works in context.
- Easy to understand.
- Ease of access.
- Sharing visualization with team.

Following is the Disadvantages of Proposed solution.

- Need access of IBM Cognos analytics on cloud.
- It can only be shared if the person has access of IBM Cognos Analytics

11. CONCLUSION

Machine learning is an important decision support tool for crop yield prediction, including supporting decisions on what crops to grow and what to do during the growing season of the crops. This dataset provides a huge amount of information on crop production in India ranging from several years. Based on the Information, machine learning algorithms like random forest and linear regression are used for prediction. Based on the performance metrics, random forest provides R squared score as 0.9978 and linear regression provides 1. Both are very good scores. Linear regression is most suited for the application.

12. FUTURE SCOPE

The model just uses parameters like crop name, season, place and area for predicting the amount of production. Apart from these, the amount of nutrients present in each land area varies within the district. So, Some more features like the amount of nutrients present in the land, good estimation of rain and sunny days can be used to give exact predictions.

13. APPENDIX

Link to the GitHub Repository:

https://github.com/IBM-EPBL/IBM-Project-1253-1658381118

Project Demo Video Link:

https://www.youtube.com/watch?v=qYDGbDcjREQ