

PROJECT REPORT

Team ID	TEAM ID - PNT2022TMID48523
Project Name	ESTIMATE CROP YIELD USING DATA ANALYTICS

1.INTRODUCTION

1.1Project Overview:

In GDP of India, more that 19% share is from Agriculture. So, it is important to analyze the Crop Production data of Indian Agriculture market. This project is aimed to create fruitful visualization using Congos Analytics on cloud for said data.

In this project various types of visualization are created to find the insights from Crop Production data of Indian market. Predictive Analytics use statistical models to analyze current and historical data for forecasting (predictions) about future or otherwise unknown events. In business, predictive analytics is used to identify risks and opportunities that aid in decision-making.

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.

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. The main challenge in using big data in agriculture is identification of effectiveness of big data analytics. Efforts are going on to understand how big data analytics can agriculture productivity. The present study gives insights on various data analytics methods applied to crop yield prediction and also signifies the important lacunae points' in the proposed area of research.

1.2 Purpose:

The purpose of this project is to know about the fundamental concepts of IBM Cognos on cloud, the working with IBM Cognos, to work with various graph and charts and to create meaningful dashboard. The analysis of data related to agriculture helps in crop yield prediction, crop health monitoring and other such related activities. In literature, there exist several studies related to the use of data analytics in the agriculture domain. Agriculture analysis is a very important aspect to crop growing. To increase quality and yields, it is crucial to understand the current nutrient levels of the soil to be able to ascertain which areas require improvement.

Data analytics can help farmers monitor the health of crops in real-time, create predictive analytics related to future yields and help farmers make resource management decisions based on proven trends. Reducing waste and improving

profits. 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.

2.LITERATURE SURVEY:

In this problem the dataset of Crop production is used. The dataset was downloaded from Kaggle using <https://www.kaggle.com/abhinand05/crop-production-in-india> link. The dataset contains the data of Agriculture production from year 1997 to 2015 in csv format. The data have 7 columns i.e. State, District, Year, Season, Area and Production. There were 24000+ data in csv. The problem is find meaningful graphs from the data. For meaningful insights we have bar graph, line-graph, map chart etc. So, to create meaningful insightsthe IBM Cognos Analytics on cloud was used to create the dashboard. It is very easy and user friendly tool by which with minimum time we can create analytics. For such analysis we can Aslo use Microsoft Excel, but when data size is large, it is difficult to handle Microsoft Excel.

2.1 References:

PAPER 1:

TITLE:Rice Crop Yield Prediction using Data Mining Techniques

AUTHORNAME: Dr. M .S, Shirdhonkar

DESCRIPTION:Discussed various data mining techniques utilized for prediction of rice crop yield for the state of Maharashtra, India. WEKA tool was applied in dataset processing

PAPER 2:

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

AUTHORNAME: Dhivya B H, Manjula R, Siva Bharathi S, Madhumathi R

DESCRIPTION:Presented a survey on the different algorithms applied in the assessment and prediction of crop yield , Discussed about the mechanism of knowledge the discovery in Agricultural data mining

PAPER 3:

TITLE:A Study on Various Data Mining Techniques for Crop Yield Prediction

AUTHORNAME: Yogesh Gandge, Sandhya

DESCRIPTION:Discussed various data mining techniques employed for predicting the crop yield and signifies the importance of accurate data extraction methods of big data analytics

PAPER 4:

TITLE:Big Data for weed control and crop protection

AUTHORNAME: F K Van Evert, S Fountas, D Jakovetic, V Crnojevic, I Travlos & C Kempenaar

DESCRIPTION:Critically discussed about the challenges faced and the profound opportunities lies in the Big Data analytics in agriculture: Outlined Big Data analytics models with numerical algorithms applied Represent the importance of reforming the mined data in the form of understandable information to the farmers. Discussed about various advances, tools and algorithms applied in transforming the data in to easily understandable information to the framers and thrown a light on success story of Netherlands in achieving the maximum crop yield and their smart forming practices.

PAPER 5:

TITLE:The Impact of Data Analytics in Crop Management based on Weather Conditions

AUTHORNAME: Swarupa Rani A

DESCRIPTION:Discussed the application of mathematical model like fuzzy logic designs in optimization of the crop yield, artificial neural networks in validation studies, genetic algorithms designs in accessing the fitness of the model applied, decision trees, and support vector machines to study soil.

PAPER 6:

TITLE:A Study on Crop Yield Forecasting Using Classification Technique

AUTHORNAME:R.Sujatha, Dr.P.Isakki Devi

DESCRIPTION:Discuss the importance of comparing previous agricultural data with present to identify optimum condition favor enhanced crop yield. Envisaged the importance of best crop selection depending on the season and the climatic factors which supports enhanced crop yield.

2.2 Problem Statement Definition

Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	A farmer	Choose a good Land	I am unawareabout area	I can't find people to guideme	Dejected

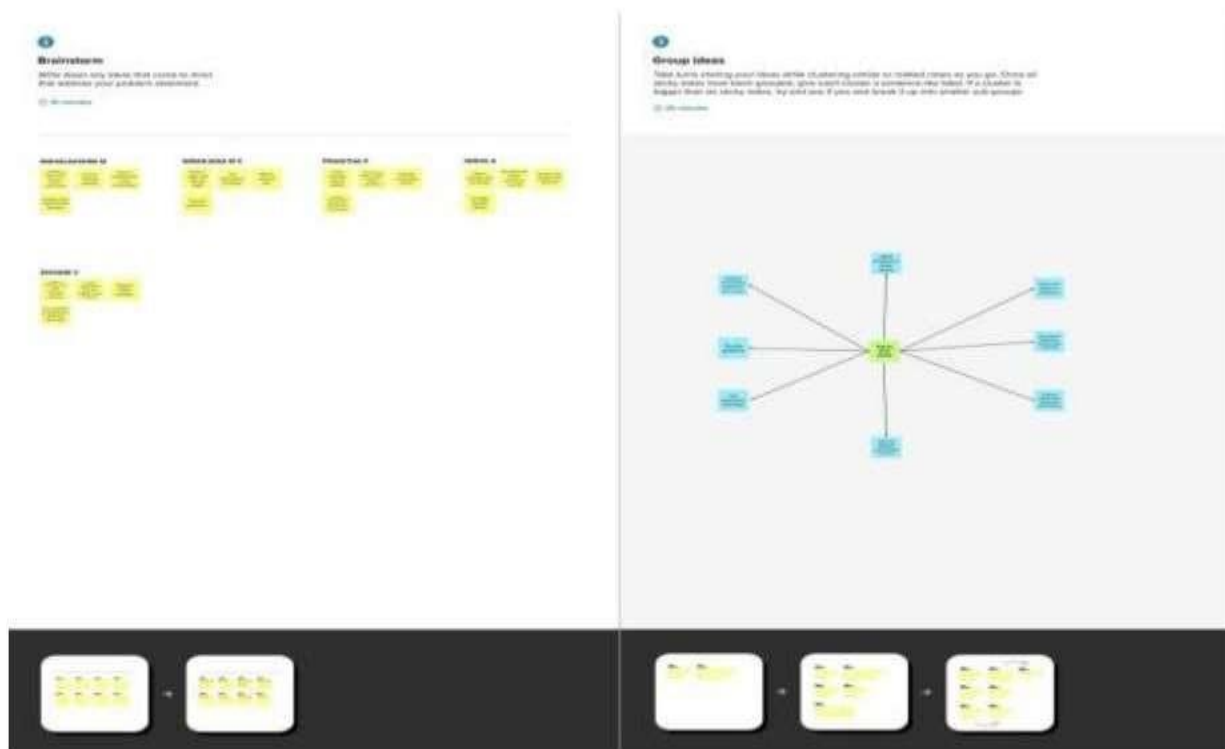
PS-2	A farmer	Buy good field for Agriculture	I am unaware about visa formalities	I can't find trusted websites	Frustrated
PS-3	A farmer	Enroll in amaster	I am unaware about sudden climate change	I can't find helping	Depressed
PS -4	A farmer	Choose good field withwell Cultivation	I am unaware about cultivation and profit	I can't find people to guideme	Anxious

3.IDEATION & PROPOSEDSOLUTION

3.1EmpathyMapcanvas



3.2 Ideation & Brainstorming

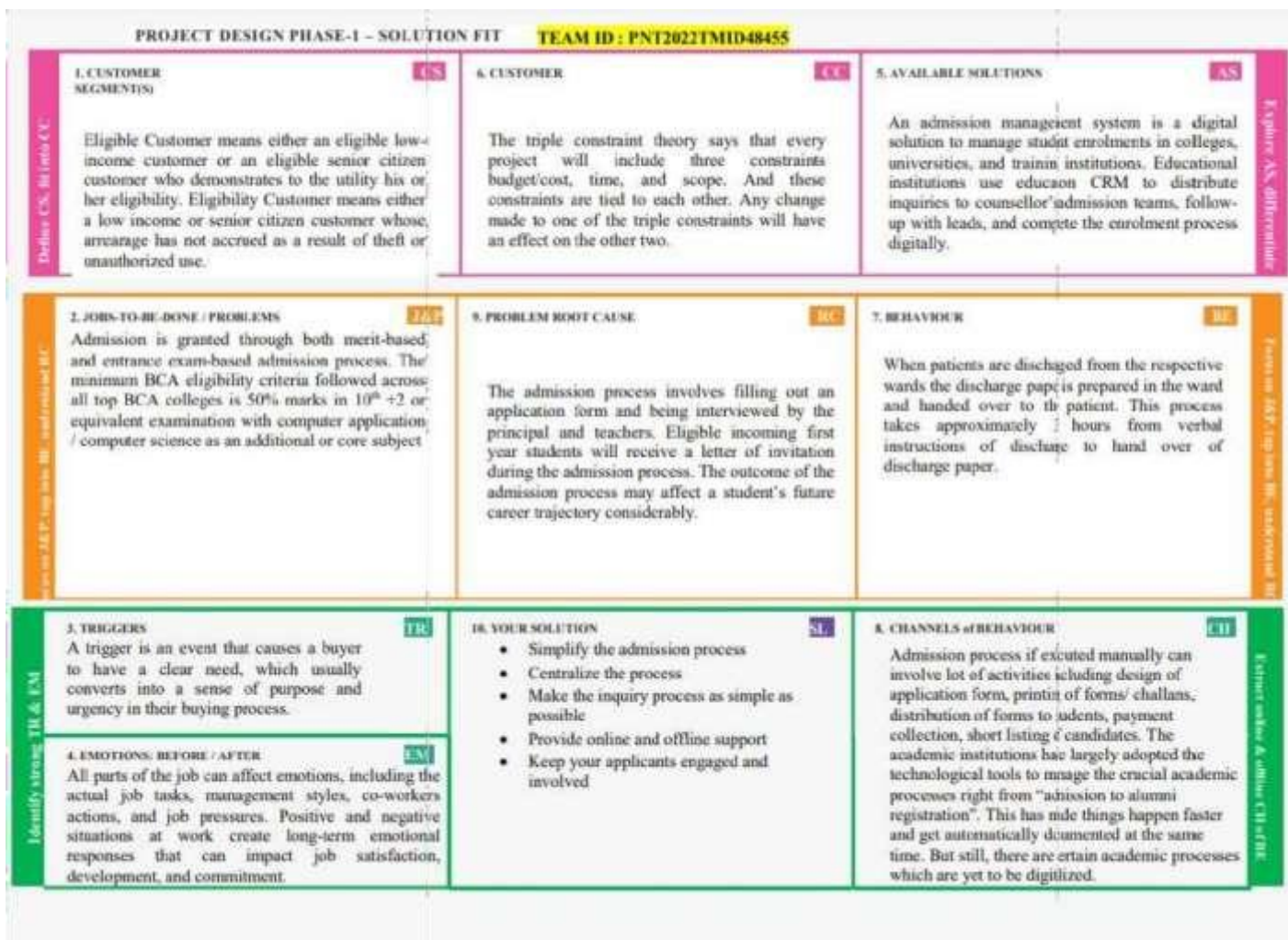


3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	To predict the crop for the farmer to who is going to cultivate in large/small field studies.
2.	Idea / Solution description	Crop yield is a standard measurement of the amount of agricultural production harvested—yield of a crop—per unit of land area. Crop yield is the measure most often used for cereal, grain, or legumes; and typically is measured in bushels, tons, or pounds per acre in the U.S.

3.	Novelty / Uniqueness	This website has to predict the universities inside India. And also give various information about the farm. Also to list the crop in the demand list.
4.	Social Impact / Customer Satisfaction	The data will reduce the panic and unawareness among farmer. It will reduce our time, travel, and costs. It will give the exact approximate prediction based on farmer.
5.	Business Model (Revenue Model)	Farm shall find the data in order to maintain it. This congos will predict and display the exact results to the farmer.
6.	Scalability of the Solution	A future update shall have crop demand.

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	Calculate crop Prediction	Enter Crop name, State Name
FR-2	Check information about the State	Visit the State of the respected crop to know more information
FR-3	Watch farntour	Check guidefor visa internet and otherprocedures
FR-4	Check financial assistance lab	Check profitand cultivation procedure

FR-5	Information about the Crop and location	Refer official State and Check the demand
------	---	---

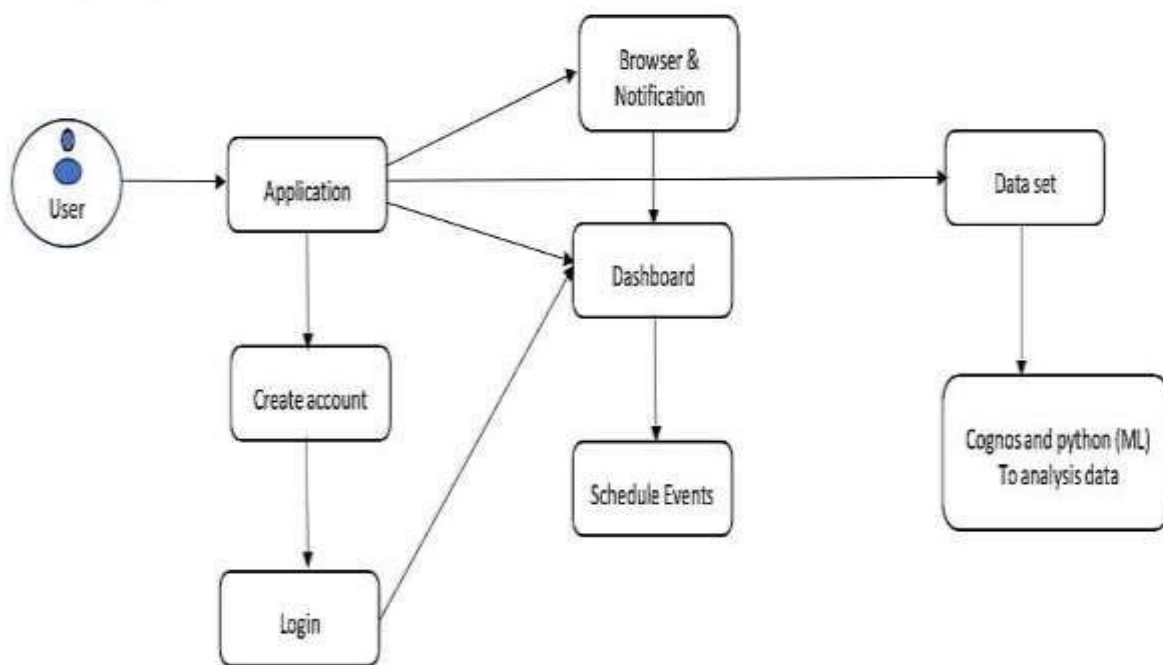
4.2 Non-Functional requirements

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	The UI/UX enhances the user experience. The entire journey of the customer throughout the application will be smooth and user-friendly approach to the user.
NFR-2	Security	It is the safest application you never used it before because it doesn't store your data.
NFR-3	Reliability	The system will give you to the most accurate and exact results.
NFR-4	Performance	Easy to for everyone to check the demand and leading crop in the available State
NFR-5	Availability	The Farm predictor will be available to users 99.9 percent of the time every month
NFR-6	Scalability	The system must be scalable enough to support more than one lacks visits at the same time while maintaining optimal performance.

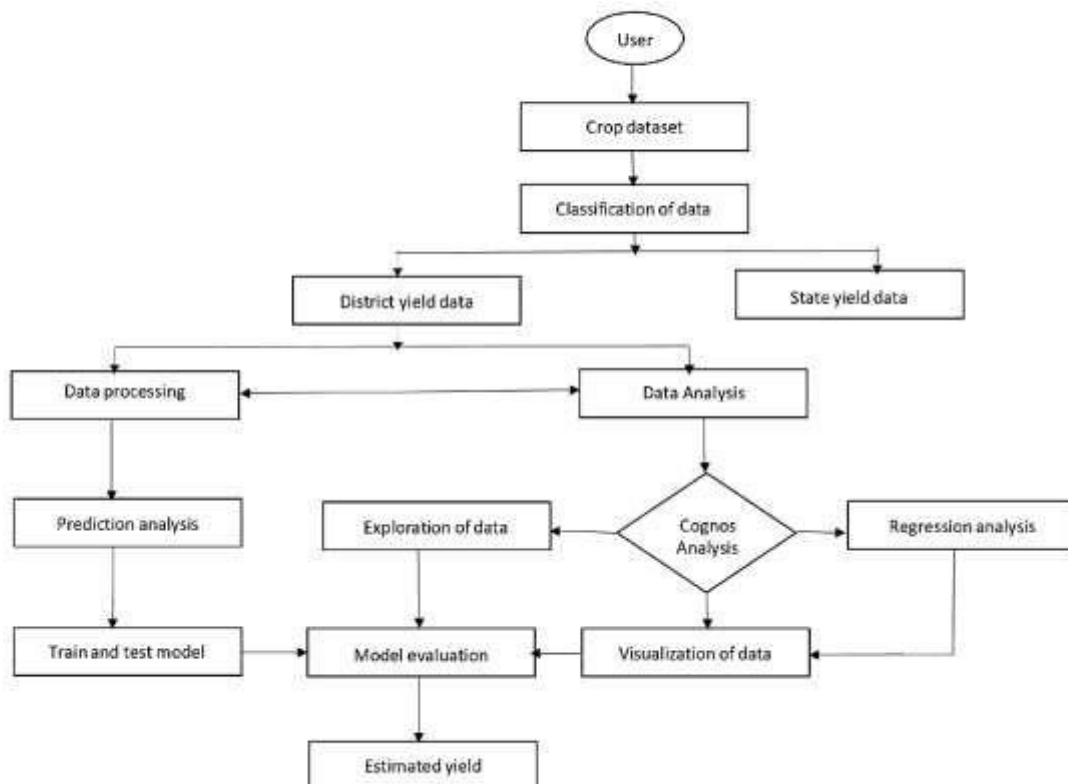
5.PROJECT DESIGN

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.



Data flow diagram:



5.2 Solution & Technical Architecture:

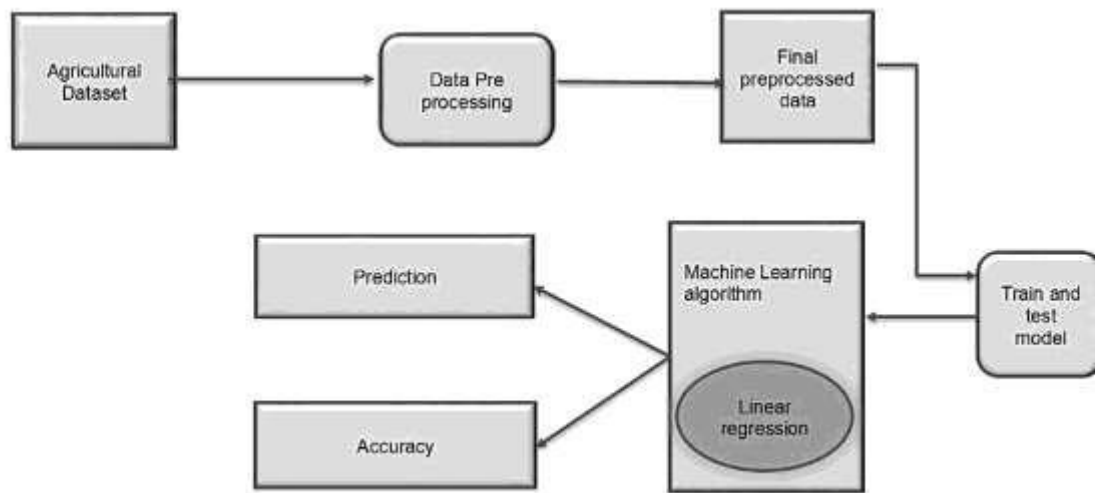


Table-1:Components&Technologies:

S.No	Component	Technology
1.	User Interface	HTML, CSS, Flask
2.	Application Logic-1	IBM Congoes
4.	Database	Dataset
5.	Cloud Database	IBM DB2,IBM Cloudant etc.
6.	File Storage	IBM BlockStorage or OtherStorage Service or Local Filesystem
7.	Machine Learning Model	Random forest, Linear Regression,etc.,.

Table-2: Application Characteristics:

S.No	Characteristics	Technology
1.	Open-Source Frameworks	Data Analytics
2.	Performance	Maximum 2mb requests per dataset

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	Login	USN-1	As a user, I can Login for the congos by entering my email, password, and confirming my password.	2	High	2
Sprint-1	with Working data set	USN-2	As a user, I will receive confirmation email once I have registered for the application	1	High	1
Sprint-1	with Working data set	USN-3	As a user, I can register for the application throughGmail	1	High	1

Sprint-1	Login	USN-4	As a user, I can log into the application by entering email & password	3	High	3
Sprint	Functional Requirement (Epic)	User Story Number	User Story/ Task	Story Points	Priority	Team Members

Sprint -2	Data visualization chart	USN-5	As a user, after logging in, I will have to update my profile by providing all the required details	5	High	5
Sprint-3	Creating Dashboard	USN-6	As a user, I will be able to view the list of Universities that the students are eligible to apply	5	Medium	5
Sprint-3	Creating Dashboard	USN-7	As a user, I will be able to view the details of Admission process like date and venue of certification verification	2	Low	2
Sprint-4	Exporting Analytics	USN-8	As a user, I will be able to view the list of courses that the students are eligible to apply	3	High	3
Sprint-1	Authentication	USN-9	As a admin, the login credential of the user is authenticated my me	2	High	2

Sprint-2	Data visualization chart	USN-10	As a admin, I can verify the user entered details	5	High	5
Sprint-3	Creating Dashboard	USN-11	As a admin, I can test the trained machine learning model by analysing the user details by machine learning algorithms like logistic regression	3	High	3
Sprint-4	Output	USN-12	As a admin, I can upload the confirmation of user for the prediction into the database.	3	High	5

6.2. Sprint 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	50 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

7.CODING

LOGIN

```

<!DOCTYPE html>
<html>
<head>
<meta name="viewport" content="width=device-width, initial-scale=1">
<title> Login Page </title>
<style>  Body
{
    font-family: Calibri, Helvetica, sans-serif; background-
color:white; background-image:
    url('https://2.bp.blogspot.com/-
pLLWS1j5PCQ/VqyQUTUqtdI/AAAAAABE64/QYlUh6421co/s1600/2de5113b6a62d0360130b90442106237_large.jp
eg');" } button {    background-color:#c3e3dc;    width: 100%;    color: purple;    padding: 15px;
margin: 10px 0px;    border: none;    cursor: pointer;
    }
    form {    border: 3px solid
    #f156189;
    }
input[type=text], input[type=password] {
width: 100%;    margin: 8px 0;
padding: 12px 20px;    display:

```

```

inline-block;      border: 2px white;
box-sizing: border-box;
}
button:hover      {
opacity: 0.7;
}
.cancelbtn {
width: auto;
padding: 10px 18px;
margin: 10px 5px;
}
.container      {
padding: 25px;
<!--      background-color: pink; -->
}
</style>
</head>
<body>
  <center> <h1>Login Form </h1> </center>
  <form>
    <div class="container">
      <label>Username : </label>
      <input type="text" placeholder="Enter Username" name="username" required>
      <label>Password : </label>
      <input type="password" placeholder="Enter Password" name="password" required>
      <button type="submit">Login</button>
      <input type="checkbox" checked="checked"> Remember me
      <button type="button" class="cancelbtn"> Cancel</button>
    <a href="#"> Forgot password? </a>

    </div>
  </form>
</body>
</html>

```

REGISTRATION

```
<!DOCTYPE html>
<html>

<head>
  <title></title>
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <link rel="stylesheet" type="text/css" href="{ {url_for('static',filename='style.css')}}">
  <link          rel="stylesheet"          href="https://cdnjs.cloudflare.com/ajax/libs/font-
awesome/4.7.0/css/fontawesome.min.css">
  <!-- jQuery library -->
  <script src="https://ajax.googleapis.com/ajax/libs/jquery/3.2.1/jquery.min.js"></script>

  <!-- Latest compiled JavaScript -->
  <script src="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.7/js/bootstrap.min.js"></script>
  <script src="https://www.google.com/recaptcha/api.js" async defer></script>
  <style      type="text/css">
body{
    margin: 10px 10px 10px 100px;    background-
color: aliceblue;
    }

    .error    {
color: red;
    }

    .fm1 {      text-
align: center;
    }

    .lb1 {      text-
align: center;
padding: 25px;
    }

    .lb2 {      margin-
left: 20px;
```

```
}
```

```
.lb3 {      margin-  
right: 35px;  
}
```

```
.container      {  
display: block;  
}      .k{  
border-radius: 15px;  
}
```

```
</style>
```

```
</head>
```

```
<body>
```

```
<?php      include  
'header.php';  
?>
```

```
<div class="heading fix">
```

```
<label class="lb1">REGISTRATION</label>
```

```
</div>
```

```
<div class="outerbox">
```

```
<div class="fixedbox">
```

```
<span class="content">
```

```
<h4>Hello, Friend!</h4>
```

```
<p>Enter your personal details and start journey with us</p>
```

```
</span>
```

```
</div>
```

```
<div class="scrollbox">
```

```
<div class="registerdonor">
```

```
<form action="process.php" method="POST" id="myform">
```

```
<div class="login">
```

```
<h3>Login Details</h3>
```

```
<table class="fm1">
```

```
<tr>
```

```
<td colspan="2">
```

```
<label class="lb1" class="username">User Name:-</label>
```



```

        <input type="text" name="user_name" required pattern="^[A-Za-z0-9._%+-@]{5,10}$"
title="Enter a username between 5 to 10 letter" autocomplete="off">
    </td>
</tr>
<tr>
    <td>
        <label class="lb1">Full Name:-</label>
        <input type="text" name="user_full_name" required pattern="[A-z ]+$"
title="Use only character & whitespace" autocomplete="off">
    </td>
</tr>
<tr>
    <td>
        <label class="lb1">Email Id:-</label>
        <input type="email" name="user_email" required
pattern="[A-Za-z0-9._%+-]+@[A-z0-9.-]+\.[a-z]{2,}$"
title="Email id is not Valid" autocomplete="off">
    </td>

```

7.RESULTS

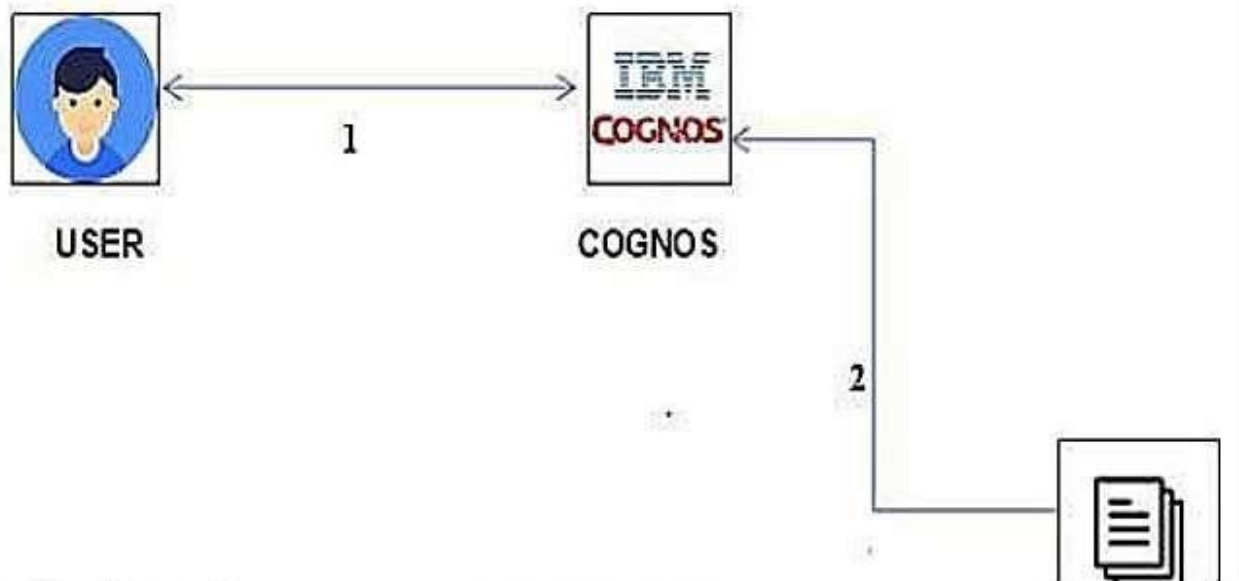


Fig. 1 Architecture Diagram

IV. RESULTS

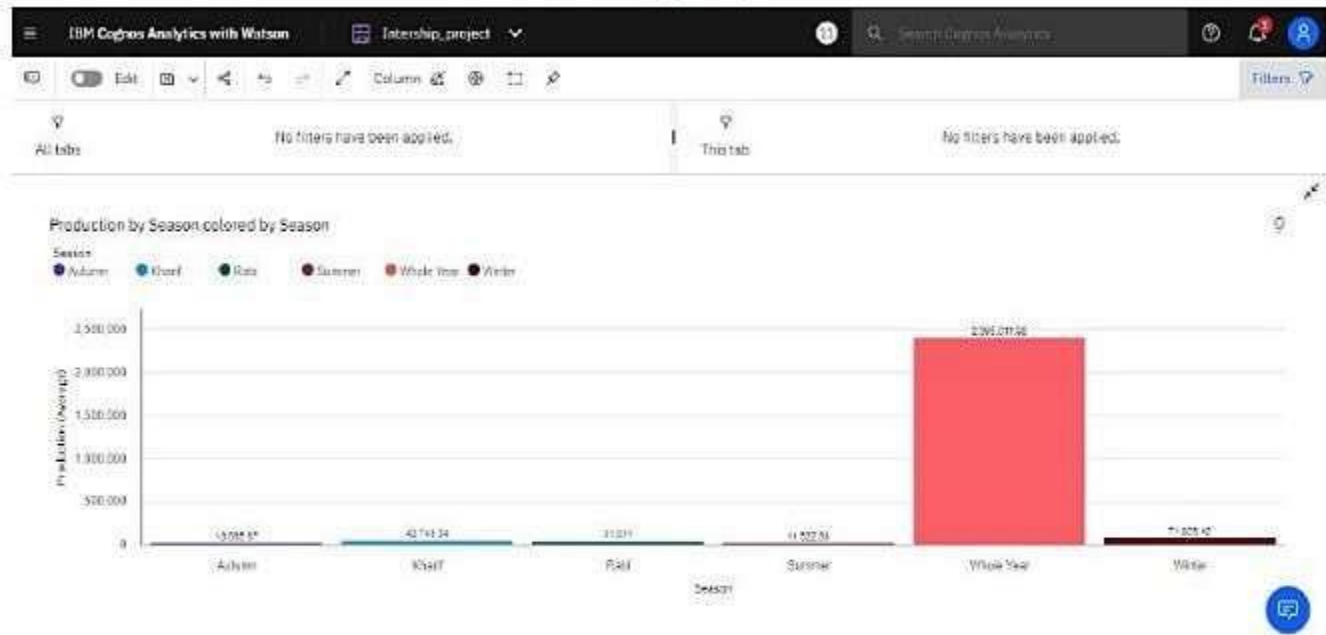


Fig. 2 Data Visualisation 1

The above figure shows the seasons with average production

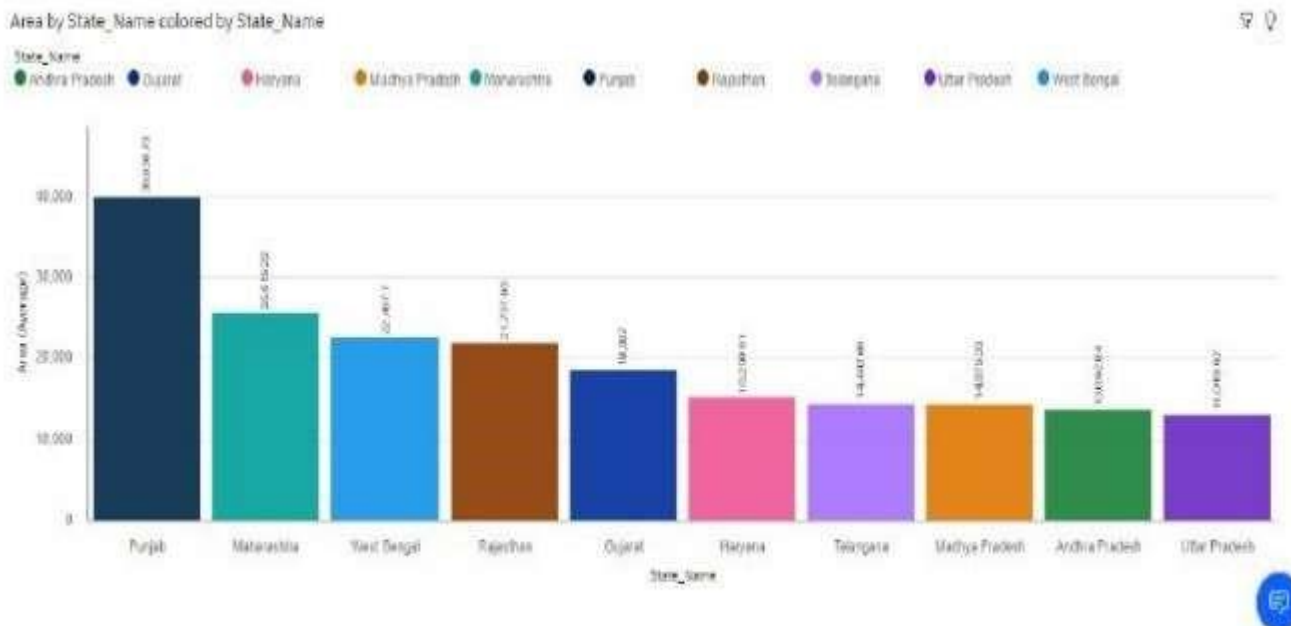


Fig. 3 Data Visualisation 2

The above figure shows the visualize top 10 states with most area

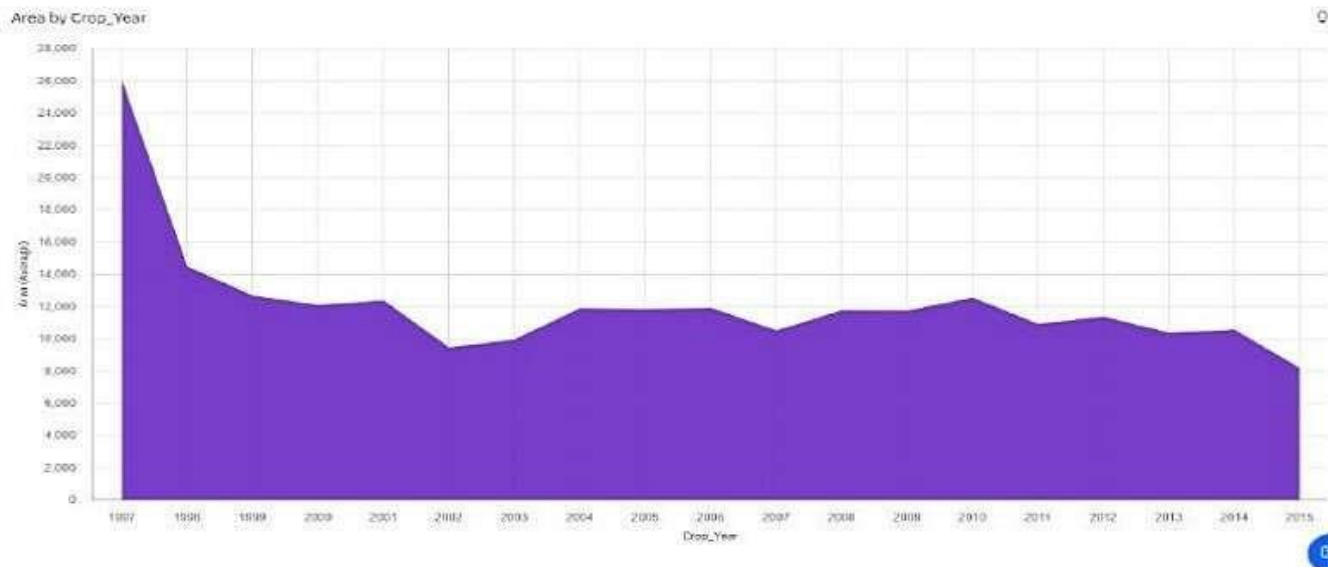


Fig. 4 Data Visualisation 3

The above figure shows the visualize with years usage of area and production

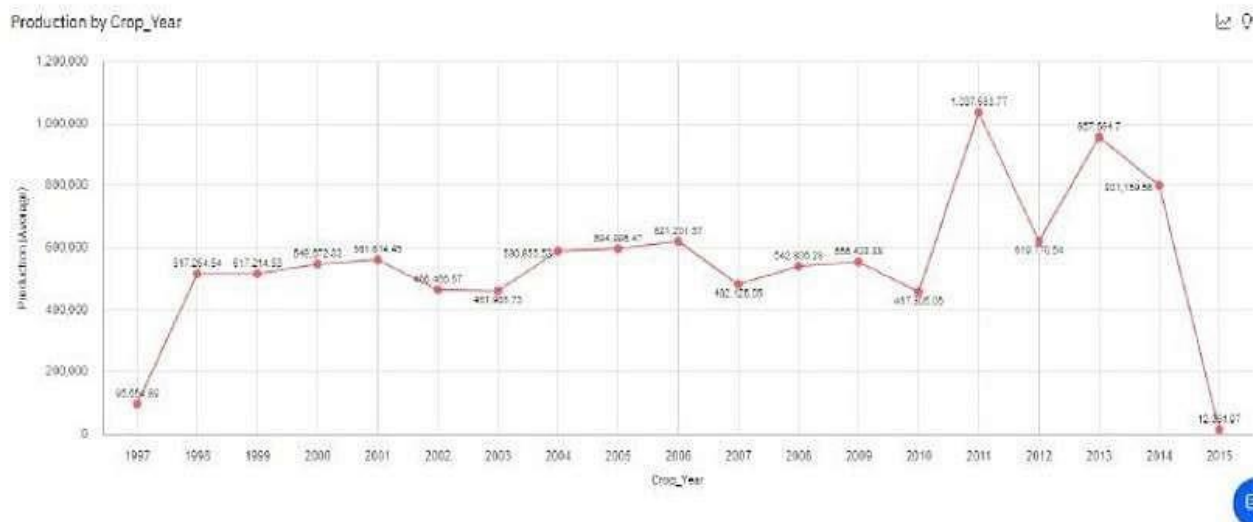
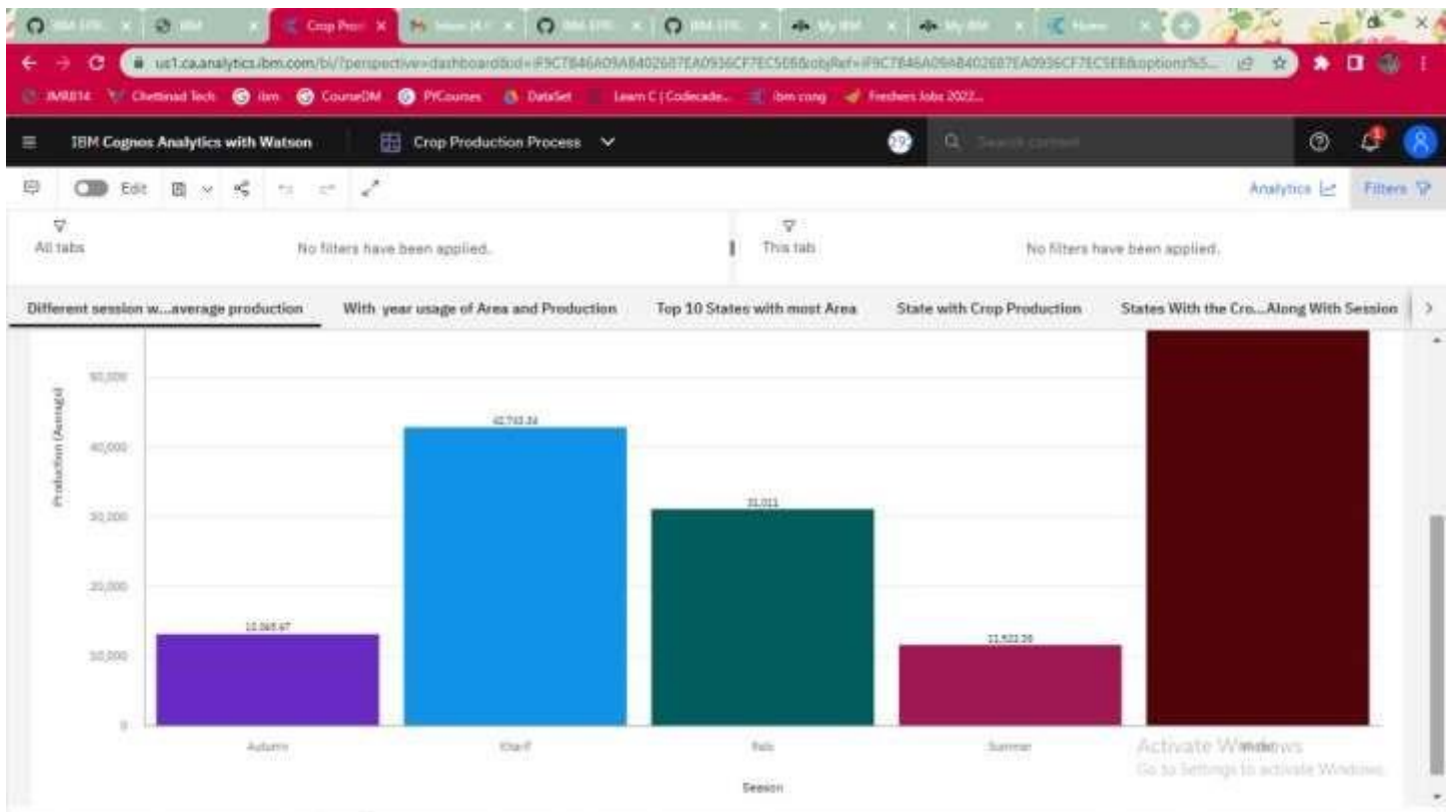


Fig.5 Data Visualisation 4

The above figure shows the visualisation contains state with crop production



IBM Cognos Analytics interface showing two tables: "State_Name and Crop" and "Season and Crop".

Crop	State_Name
Arhar/Tur	Odisha
Bajra	Gujarat
Bajra	Karnataka
Bajra	Puducherry
Banana	Manipur
Banana	Uttar Pradesh
Brinjal	Puducherry
Cotton(Int)	Karnataka
Cotton(Int)	Puducherry
Cowpea(Lobia)	Karnataka
Dry chilies	Karnataka
Dry chilies	Manipur
Dry chilies	Puducherry
Dry ginger	Manipur
Dry ginger	Gujarat
Dry ginger	Karnataka

Crop	Season
Apple	Whole Year
Arcanaut (Processed)	Whole Year
Arcanaut	Kharif
Arcanaut	Rabi
Arcanaut	Whole Year
Arhar/Tur	Autumn
Arhar/Tur	Kharif
Arhar/Tur	Rabi
Arhar/Tur	Summer
Arhar/Tur	Whole Year
Arhar/Tur	Winter
Ash Gourd	Whole Year
Atacnut (Raw)	Whole Year
Bajra	Kharif
Bajra	Rabi
Bajra	Summer

IBM Cognos Analytics with Watson interface showing the "Selected sources" panel and a visualization area.

Selected sources:

- Navigation paths
 - crop_production.csv
 - State_Name
 - District_Name
 - Crop_Year
 - Season
 - Crop
 - Area
 - Production

Visualization area:

Tab 1

No visualization selected

Select a visualization to manage the field settings.

Activate Windows

us1.ca.analytics.ibm.com/bl/perspective=dashboard&id=99CT646A09A8402687EAD936CF7EC5D5&objRef=99CT646A09A8402687EAD936CF7EC5D5&...

IBM Cognos Analytics with Watson Crop Production Process

Assistant

crop_production.csv

11/18/2022 | 9:39 AM
Connected to crop_production.csv

@Cognos

How can I help you?

All tabs No filters have been applied. This tab

w...average production With year usage of Area and Production Top 10 States with most Area

State_Name and Crop

Crop	State_Name
Arhar/Tur	Odisha
	Gujarat
Bajra	Karnataka
	Puducherry
Banana	Manipur
Brinjal	Uttar Pradesh
	Puducherry
Cotton(jint)	Karnataka
	Puducherry

Season and Crop

Crop	Season
Apple	Whole Year
Arcanot (Process...	Whole Year
	Kharif
Arcanot	Rabi
	Whole Year
	Autumn
	Kharif
	Rabi
Arhar/Tur	

24°C Mostly cloudy 18-11-2022

us1.ca.analytics.ibm.com/bl/perspective=ca-modeler&id=3370522320_99873645e14cfaad74913504890418_sessionTemp&objRef=3370522320_99873645e14cfaad74913504890418_sessionTemp&...

IBM Cognos Analytics with Watson * New data module

Data module

Search

New data module

Navigation paths

crop_production.csv

Row Id

State_Name

District_Name

Crop_Year

Season

Crop

Area

Production

Grid Relationships Custom tables

Row Id	State_Name	District_Name	Crop_Year	Season	Crop
1	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Arecanut
2	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pul
3	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice
4	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana
5	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashewnut
6	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Coconut
7	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Dry ginger
8	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sugarcane
9	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sweet potato
10	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Tapioca
11	Andaman and Nicobar Islands	NICOBARS	2001		Arecanut

24°C Mostly cloudy 18-11-2022

IBM Cognos Analytics with Watson

New data module

Search

New data module

- Navigation paths
- crop_production.csv
 - Row Id
 - State_Name
 - District_Name
 - Crop_Year
 - Season
 - Crop
 - Area
 - Production

Row Id	State_Name	District_Name	Crop_Year	Season	Crop
1	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Areca nut
2	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pul
3	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice
4	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana
5	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashew nut
6	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Coconut
7	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Dry ginger
8	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sugarcane
9	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sweet potato
10	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Tapioca
11	Andaman and Nicobar Islands	NICOBARS	2001		Areca nut

Activate Windows
Go to Settings to activate Windows.

24°C Mostly cloudy

09:36 PM 16-11-2022

File | C:\Users\MY\Downloads>Login%20form.html

Login Form

Username :

Password :

Login

☒ Remember me

24°C Partly cloudy

01:13 AM 16-11-2022

REGISTRATION

Hello, Friend!

Enter your personal details and start journey with us

Login Details

User Name:-

Full Name:-

Email Id:-

Password:-

Confirm Password:-

Contact Details

Mobile Number:-

Pincode

Address:-

State:-

Activate Windows
Go to Settings to activate Windows.

8.ADVANTAGES & DISADVANTAGES

ADVANTAGES

- increase the yield from existing farmlands
- Big data provides farmers granular data on rainfall patterns, water cycles, fertilizer requirements, and more. This enables them to make smart decisions, such as what crops to plant for better profitability and when to harvest.
- In the pre-processing step dataset will be divided into training dataset and testing dataset. This is the important step while creating model.
- It is implemented in congos with and dataset packages, which are open source and freely available.
- It uses many algorithms like random forest,K-nearest neighbour, multilayer preceptron.
- The simplest way to determinewhether a person is eligible for a university or college as well as being completely objective and transparent.

DISADVANTAGES

- It is used to display the result as a dotted graph.
- It is to predict the approximate value.
- Contrarily, linear regression presumes that the relationship between the dependent and independent variables is linear. This implies that it considers their relationship to be linear. The independence of the qualities is assumed.

It is not always true that the conditional independence assumption is true.

9.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 yield prediction etc. In this survey, the specific activity, crop yield prediction has been surveyed and the major trends have been identified. The rice crop yield prediction has been done in the state of Maharashtra using data mining techniques in one of the works.

The analysis has been done using machine learning framework WEKA. In the work carried out in, various algorithms applied in the assessment of crop yield and mechanism for knowledge discovery have been discussed. The challenges and opportunities in the field of Big Data analytics in agriculture have been discussed in with a case study of Netherlands.

10.FUTURE SCOPE

The future scope of this project is very broad. Few of them are:

- This can be implemented in less time for creating graph.
- This can be accessed anytime anywhere, since it is a web application provided only an internet connection.
- The user had not need to travel a long distance for the crop and his/her time is also saved as a result of this automated system.

GitHub Link:

<https://github.com/IBM-EPBL/IBM-Project-36340-1660294318>

Video Demo Link:

https://youtu.be/t_w1M-A8CPo