Er. PERUMAL MANIMEKALAI COLLEGE OF ENGINEERING
ACCREDITED BY NBA & NAAC WITH 'A' GRADE
Koneripalli, HOSUR - 635 117.

# Exploratory Analysis of Rain Fall Data in India for Agriculture

A PROJECT REPORT
Submitted by

KRUTHIKA.S.J (610819205023)
SANGEETHA.R (610819205044)
POOJA.L (610819205037)
SUMA.S (610819205051)
TEAM ID-(PNT2022TMID30080)

MENTOR: MRS.R.VASUGI(M.TECH)

in partial fulfillment for the award of the degree
of
**BACHERLOR OF TECHNOLOGY**
in
**INFORMATION TECHNOLOGY**

Er.PERUMAL MANIMEKALAI COLLEGE OF ENGINEERING,HOSUR

ANNA UNIVERSITY::CHENNAI 600 025

NOVEMBER 2022

# CONTENTS

# 1.INTRODUCTION

Rainfall pattern means the distribution of rain geographically, temporally, and seasonally. The tropics receive more rainfall than deserts. Cooler places receive no rainfall, as it is converted to snow before it falls to the ground. Rainfall happens more in a particular time of a year, during a rainy season. In other seasons, rainfall is scant. Therefore, agriculture, worldwide, is planned based on rainfall's natural pattern. Water reservoirs, irrigation networks, and urban water supply systems are designed based on the average annual rainfall. If it rains a lot on a continuous basis for a longer time, there is a possibility of flood and subsequent disaster to the infrastructure. No rainfall or little rainfall for a longer period in an inhabited area could lead to drought and famine.

There are two approaches to predict rainfall. They are Empirical methods and dynamical methods. The empirical approach is based on analysis of historical data of rainfall and its relationship to a variety of atmospheric and oceanic variables over different parts of the world. The most widely used empirical approaches used for climate prediction are regression, ANN, fuzzy logic and group method of data hand.

This includes data for agriculture, infrastructure, technology and so on. And so we can look at some of this data to see if we can glean something meaningful. we look at the agricultural food production for various grains, cereals and oilseeds from the year 2001 to 2017. We have also obtained data for annual rainfall in India for these years as a separate dataset and our goal is to see how strongly agricultural food production depends on annual rainfall.

Although the subject area of climate change is vast, the changing pattern of rainfall is a topic within this field that deserves urgent and systematic attention, since it affects both the availability of freshwater and food production.Based on experimentation at New Delhi, India, Aggarwal has reported that a 1°C rise in temperature throughout the growing period will reduce wheat production by 5 million tonnes. The global average precipitation is projected to increase, but both increases and decreases are expected at the regional and continental scales.Higher or lower rainfall, or changes in its spatial and seasonal distribution would influence the spatial and temporal distribution of runoff, soil moisture and groundwater reserves, and would affect the frequency of droughts and floods.

The Indian climate is dominated by the southwest monsoon. About 80% of the rainfall in India occurs during the four monsoon months (June–September) with large spatial and temporal variations over the country. Such a heavy concentration of rainfall results in a scarcity of water in many parts of the country during the non-monsoon period. Therefore, for India, where agriculture has a significant influence on both the economy and livelihood, the availability of adequate water for irrigation under changed climatic scenarios is very important. The agricultural output is primarily governed by timely availability of water. In future, population growth along with a higher demand for water for irrigation and industries will put more pressure on water resources.

With the growing recognition of the possibility of adverse impacts of global climate change on water resources, an assessment of future water availability at various spatial and temporal scales is needed. It is expected that the response of hydrological systems, erosion processes and sedimentation could significantly alter due to climate change. An understanding of the hydrological response of a river basin under changed climatic conditions would help solve problems associated with floods, droughts and allocation of water for agriculture, industry, hydropower generation, domestic and industrial use. Scenarios of changes in runoff and its distribution depend on the future climate scenarios.

In India, attempts have been made in the past to determine trends in the rainfall at national and regional scales. Most of the rainfall studies were confined to the analysis of annual and seasonal series for individual or groups of stations. In the present study, a much wider view has been taken, and changes in rainfall have been studied on seasonal and annual scales for 30 sub-divisions, and five main regions. Intra-seasonal variability in rainfall has also been studied by analysing the trends in monthly rainfall. Further, the time series of rainfall data used in this study spans more than 100 years. Thus, the present analysis is a significant improvement over the studies carried out previously.

### 1.1.PROJECT OVERVIEW

A change detection study using monthly rainfall data for 306 stations distributed across India was attempted by Rupa Kumar *et al*. (1992). They showed that areas of the northeast peninsula, northeast India and northwest peninsula experienced a decreasing trend in summer monsoon rainfall. A widespread increasing trend in monsoon rainfall over the west coast, central peninsula and northwest India was also reported. The decreasing trend ranged between −6 and −8% of the normal per 100 years, while the

increasing trend was about 10–12% of the normal per 100 years. Srivastava *et al*. (1998) supported the existence of a definite trend in rainfall over smaller spatial scale. Sinha Ray & De (2003) concluded that all-India rainfall and surface pressure shows no significant trend, except for some periodic behaviour. According to Sinha Ray & Srivastava (1999), the frequency of heavy rainfall events during the southwest monsoon has shown an increasing trend over certain parts of the country, whereas a decreasing trend has been observed during winter, pre-monsoon and post-monsoon seasons. These authors tried to attribute this variation to dynamic and anthropogenic causes. The inter-annual and decadal variability in summer monsoon rainfall over India was examined by Kripalani *et al*. (2003) by using observed data for a 131-year period (1971–2001). They found random fluctuations in annual rainfall and distinct alternate epochs (lasting approximately three decades) of above- and below-normal rainfall for decadal rainfall. They also concluded that this inter-annual and decadal variability appears to have no relationship to global warming. Analysis of rainfall data for the period 1871–2002 indicated a decreasing trend in monsoon rainfall and an increasing trend in the pre-monsoon and post-monsoon seasons.

Based on meteorological considerations, India has been divided into 36 meteorological sub-divisions (34 on the main land and two on islands). The sub-divisions on the mainland are shown.In this study, rainfall over the whole of India (except the Hilly region and islands) was considered. The geographical area of the sub-divisions considered in this study is 2.88 × 106 km2 (cf. 3.29 × 106 km2, being the whole of India, excluding islands).

Sub-divisional monthly rainfall data of India prepared by the Indian Institute of Tropical Meteorology were used in this study. A network of 306 stations (one representative station per district) over 30 meteorological sub-divisions was used to prepare the sub-divisional data.The monthly (January–December) area-weighted rainfall series for each of the 30 meteorological sub-divisions were prepared by assigning the district area as the weight for each raingauge station in that sub-division. The station rainfall data were obtained from the India Meteorological Department (IMD). Before releasing the data, the IMD carries out quality checks to ensure that error-free data are used in analysis and design. Thus the quality of this data set is very good and it is one of the most reliable long series of data. The monthly data were available for 135 years (1871–2005) for 30 sub-divisions.

Rainfall data of six meteorological sub-divisions, namely Jammu and Kashmir, Uttaranchal, Himachal Pradesh, Arunachal Pradesh, Lakshadweep and Andaman & Nicobar islands, were not available. As may be seen from the area of studied sub-divisions varies from a minimum of 18,817 km2 (Coastal Karnataka) to a maximum of 1,95,086 km2 (West Rajasthan) with the number of rainfall stations varying from two (Coastal Karnataka) to 26 (East Uttar Pradesh).

To investigate the changes in rainfall for different seasons, a year was divided into four seasons: winter (December–February), pre-monsoon (March–May), monsoon (June–September), and post-monsoon (October–November). Rainfall analysis was carried out for all the seasons as well as the whole year separately. Note that the post-monsoon season contains only two months, while the monsoon season has four months. For the trend analysis, monthly rainfall series were used to form seasonal and annual series of these variables. Basic statistics, such as minimum, maximum, mean and coefficient of variation (CV) of annual rainfall of different data sets are given in the temporal variation of annual rainfall for the five regions and the entire study area. Note that the annual rainfall of the North West India region is about half that of the other regions and its CV is about twice that for the other regions.Low rainfall coupled with a high variation makes this region highly vulnerable to climate change.

## 1.2.PURPOSE

Rainfall forecasting is very important because heavy and irregular rainfall can have many impacts like destruction of crops and farms, damage of property so a better forecasting model is essential for an early warning that can minimize risks to life and property and also managing the agricultural farms in better way. This prediction mainly helps farmers and also water resources can be utilized efficiently. Rainfall prediction is a challenging task and the results should be accurate. There are many hardware devices for predicting rainfall by using the weather conditions like temperature, humidity, pressure. These traditional methods cannot work in an efficient way so by using machine learning techniques we can produce accurate results. We can just do it by having the historical data analysis of rainfall and can predict the rainfall for future seasons. We can apply many techniques like classification, regression according to the requirements and also we can calculate the error between the actual and prediction and also the accuracy. Different techniques produce different accuracies so it is important to choose the right algorithm and model it according to the requirements.

The predictive model is used to prediction of the precipitation. The first step is converting data in to the correct format to conduct experiments then make a good

analysis of data and observe variation in the patterns of rainfall. We predict the rainfall by separating the dataset into training set and testing set then we apply different machine learning approaches (MLR,SVR) and statistical techniques and compare and draw analysis over various approaches used. With the help of numerous approaches we attempt to minimize the error.

The dataset consists of the measurement of rainfall from year 1901-2015 for each state.

• Data consists of 19 attributes (individual months, annual, and combinations of 3 consecutive months) for 36 sub divisions.
• The data is available only from 1950 to 2015 for some of the subdivisions.
• The attributes are the amount of rainfall measured in cm.

# 2.LITERATURE SURVEY

In India, Agriculture contributes major role to Indian economy. For agriculture,

Rainfall is important but during these days' rainfall prediction has become a major challenging problem. Good prediction of rainfall provides knowledge and know in advance to take precautions and have better strategy about theirs crops. Global warming is also having severe effect on nature as well as mankind and it accelerates the change in climatic conditions. Because of its air is getting warmer and level of ocean is rising, leads to flood and cultivated field is changing into drought. Due to adverse climatic change leads to unseasonable and unreasonable amount of rainfall. To predict Rainfall is one of the best techniques to know about rainfall and climate. The main aim of this study revolves around providing correct climate description to the clients from various perspectives like agriculture, researchers, generation of power etc. to grasp the need of transformation in climate and its parameters like temperature, humidity, precipitation, wind speed that eventually directs to projection of rainfall. Rainfall also depends on geographic locations hence is an arduous task to predict. Machine Learning is the evolving subset of an AI, that helps in predicting the rainfall. In this research paper, we will be using UCI repository dataset with multiple attributes for predicting the rainfall. The main aim of this study is to develop the rainfall prediction system and predict the rainfall with better accuracy with the use of Machine Learning classification algorithms.

Rainfall prediction is the one of the important techniques to predict the climatic conditions in any country. This paper proposes a rainfall prediction model using Multiple Linear Regression (MLR) for Indian dataset. The data taken from 1901 to 2015 monthly wise. The input data is having multiple meteorological parameters and to predict the rainfall in more precise. The Mean Square Error (MSE), accuracy, correlation are the parameters used to validate the proposed model.

## 2.1.EXISTING PROBLEM

With the issues with our original dataset, we learned many things considering all the preprocessing steps that we carried to rectify them. The first important thing we learned is the importance of knowing your data. While imputing the missing value, we grouped two other features and calculated the mean instead ofdirectly calculating the

mean for all the instances. This way our imputed values were closer to the correct information. Another thing we learned is about the leaky features. While exploring our data, we came to that one of our feature(RiskMM) was used for generating the target variable and hence it made no sense to use this feature for predictions.We learned about the curse of dimensionality while dealing with categorical variables which we solved using feature hashing. We also learned two techniques for performing feature selection - univariate selection and correlation heat map.We also explore under sampling and oversampling techniques while handling theclass imbalance problem.

**2.2.REFERENCE**


**[1]** Analysis of Rainfall Trends and Its Relationship with SST Signals in the *Lake Tana Basin*: ***Tesfay Mekonnen Weldgerima, Belay Simane Birhanu (16 May 2018)***

**[2]** Long term Analysis of Rainfall data of Fatehabad Districts of Haryana: ***Arbinda Sharma, Mohit Kumar***

**[3]**. Rainfall variability and its implications for agricultural Production in Gedarif State, Eastern Sudan: ***Lotfie A. Yousif, Faisal M. EI-Hag (27 July 2018)***

**[4]**. Analysis of rainfall data for storage and irrigation planning in humid south-eastern plain (Rajasthan): ***Bhim Singh, C.K. Arya, K.K. Maurya (25 May 2014)***

**[5]**. Frequency analysis of extream rainfall events:***Salaheddine EI Adlouni***

**[6]**. A Data – Driven Approach for Accurate Rainfall Prediction: ***Shilpa Manandhar,Soumyabrata***

***[7]*** **.**Seasonal predictability of Ethiopian Kiremt rainfall and forecast skill of ECMWF's SEAS5 model. Climate Dynamics. 2021:**Ehsan MA.**

***[8]***. Rainfall Prediction using Data Mining Techniques: A Systematic Literature Review. *Int. J. Adv. Comput. Sci. Appl.* 2018;**9**:143–150. doi: 10.14569/IJACSA.2018.090518: **Aftab S., Ahmad M., Hameed N., Salman M., Ali I., Nawaz Z**

## 2.3.PROBLEM STATEMENT DEFINATION

Climate is a important aspect of human life. So, the Prediction should accurate as much as possible. In this paper we try to deal with the prediction of the rainfall which is also a major aspect of human life and which provide the major resource of human life which is Fresh Water. Fresh water is always a crucial resource of human survival – not only for the drinking purposes but also for farming, washing and many other purposes.

Making a good prediction of climate is always a major task now a day because of the climate change.

Now climate change is the biggest issue all over the world. Peoples are working on to detect the patterns in climate change as it affects the economy in production to infrastructure. So as in rainfall also making prediction of rainfall is a challenging task with a good accuracy rate. Making prediction on rainfall cannot be done by the traditional way, so scientist is using machine learning and deep learning to find out the pattern for rainfall prediction.

A bad rainfall prediction can affect the agriculture mostly framers as their whole crop is depend on the rainfall and agriculture is always an important part of every economy. So, making an accurate prediction of the rainfall somewhat good. There are number of techniques are used of machine learning but accuracy is always a matter of concern in prediction made in rainfall. There are number of causes made by rainfall affecting the world ex. Drought, Flood and intense summer heat etc. And it will also affect water resources around the world. Our major concern is the major downfall to the rainfall on yearly in the major downfall to the yearly rainfall in centimetre.

The weather has a significant impact on the agricultural industry and because of that, being able to predict it helps farmers in their day-to-day decisions such as how to plan efficiently, minimize costs and maximize yields.

A major agricultural company needs you to help them maximize growth efficiency, save resources and optimize their production.Rainfall, despite being such an important natural resource, is difficult to predict when even the slightest change in factors affecting it comes into play. Thus, we are aiming to create a system which can accurately predict the rainfall, for such a multitude of factors.

| | |
|---|---|
| I am | • Farmers who cultivates multiple crops. |
| I'm trying to | • Cultivate healthy crops to meet the raising demand for food |
| but | • Due to heavy rains crops may get affected which might result in infected crops |
| because | • I am not aware of possibilties of rainfall |
| which makes me feel | • sad,my efforts gets wasted and it affects economically too. |

# 3.IDEATION AND PROPOSED SOLUTION

## 3.1.EMPATHY MAP CANVAS

Empathy maps are an efficient tool used by designers to not only understand user behavior, but also visually communicate those findings to colleagues, uniting the team under one shared understanding of the user. Originally invented by Dave Gray at Xplane, the empathy map was made in an attempt to limit miscommunication and misunderstanding about target audiences, including customers and users.

**WHAT DO THEY THINK AND FEEL?**

- How measurement of rainfall is calculate?
- Whether this estimation will works in summer?
- Possibilities of raining?
- Does this guesstimate helps in saving water?
- Can suitable harvesting be done?
- Is the data set area accurate?

**WHAT DO THEY SEE?**

- Closer mounting plants
- Cost will be Worthful
- Expected proft in selling harvested crops
- Impact of climate change and environment

**WHAT DO THEY HEAR?**

- How is it possible used in agriculture?
- How the relevant data set are created?
- Enrichment in quality of crops
- When and Where need to be used?
- What is innovative in this?

**WHAT DO THEY SEE AND DO?**

- Producing Healthy crops
- Lessen the crop damage
- Maintaining farm
- Makes research about this analysis
- increasing annum income for farmers


## 3.2.IDEATION AND BRAINSTROMING

### Ideation:

Ideation is the process where you generate ideas and solutions through sessions such as,Sketching,Prototyping ,Brainstorming, Brainwriting, Worst Possible Idea, and a wealth of other ideation techniques. Ideation is also the third stage in the Design Thinking process. Although many people might have experienced a “brainstorming” session before, it is not easy to facilitate a truly fruitful ideation session. In this article, we’ll teach you some processes and guidelines which will help you facilitate and prepare for productive, effective, innovative and fun ideation sessions.

Ideation is often the most exciting stage in a Design Thinking project, because during Ideation, the aim is to generate a large quantity of ideas that the team can then filter and cut down into the best, most practical or most innovative ones in order to inspire new and better design solutions and products.

**BRAINSTORMING:**

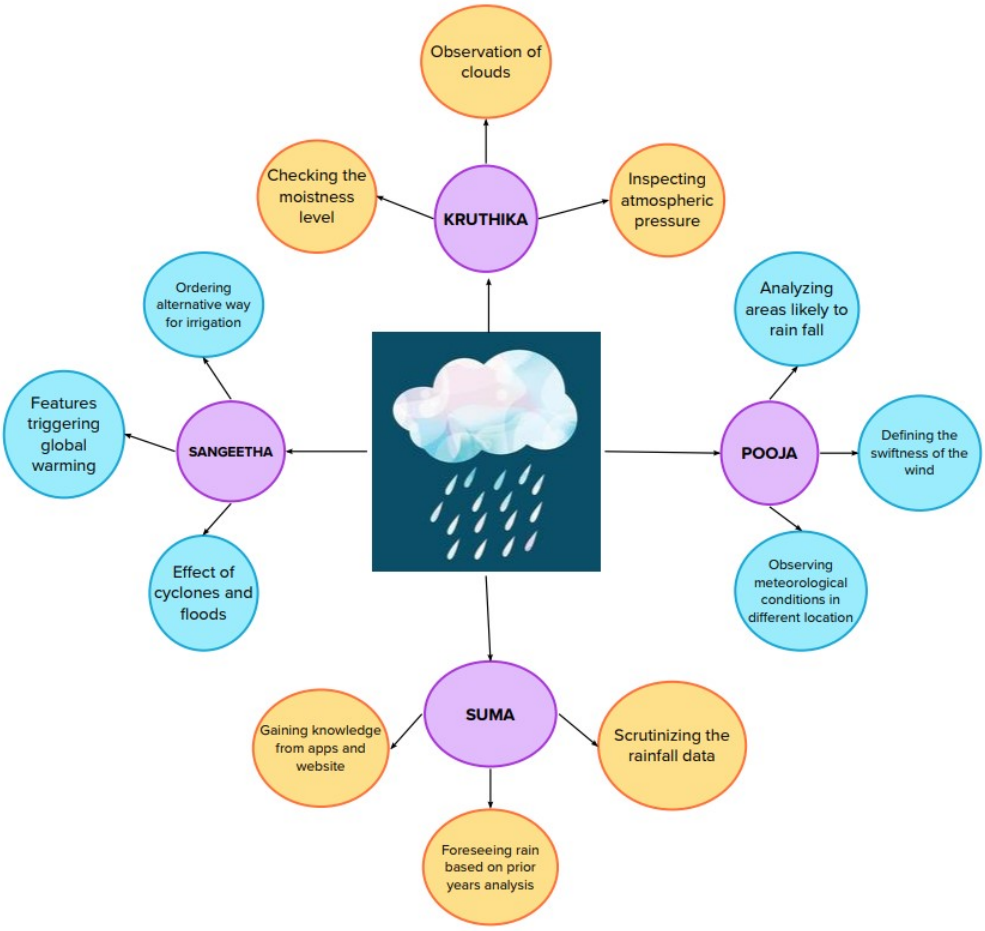
- Bainstorming is a method of generating ideas and sharing knowledge to solve a particular commercial or technical problem, in which participants are encouraged to think without interruption. Brainstorming is a group activity where each participant shares their ideas as soon as they come to mind.
- Brainstorming combines a relaxed, informal approach to problem solving with lateral thinking. It encourages people to come up with thoughts and ideas.
- Some of these ideas can be crafted into original, creative solutions to a problem, while others can spark even more ideas.
- Brainstorming is a strategy used to generate a number of ideas to help solve a particular problem.The concept of idea building is sharing ideas, which triggers new ideas, which creates a chain of new thoughts.

Some basic ideas:

- Analyzing areas likely to rainfall
- Inspecting atmospheric pressure
- Effect of cyclones and food
- Gaining knowledge from website

Prioritize:

- Observing meteorological conditions in different location
- Foreseeing rain based on prior years analysis
- Checking the moistness level
- Scrutinizing the rainfall data




## 3.3.PROPOSED SOLUTION

- Rainfall is very important because heavy and irregular rainfall can have many impacts like the destruction of crops and farms.
- Rainfed crops are prone to breaks in the monsoon during the crop growth due to water stress.
- Farmers need to satisfy the rising demand for more food of higher quality.

Solution description:

- Water resources can be fully untilised through interstate coordination on water management, water can easily be delivered to the areas where it is most required.

- Helps in producing healthy crops and good fields.
- Provides extra support to maintain the agriculture.

Uniqueness:

- This application is useful for the beginners in agriculture.
- Seed maturity selection features are available.

Social Impact:

- Different types of crops can be planted for good health.
- It helps in making of decisions regarding the cropping pattern and sowing rate.

Business Model:

- Indian economy is vitally linked with the monsoon because of its water resources.
- The growth of economy is stimulated by higher production of crops which makes profit in trading.

Scalability of the Solution:

- A good balance of rain and proper irrigation can lead to faster growing plants and yields better crops.

## 3.4.PROBLEM SOLUTION FIT

The Problem-Solution Fit simply means that you have found a problem with your customer and that the solution you have realized for it actually solves the customer's problem.

Problem-solution fit is a term used to describe the point validating that the base problem resulting in a business idea really exists and the proposed solution actually solves that problem.

CUSTOMER SEGMENT:

- Farmers are the customers who cultivate different types of crops depending on the rainfall.
- Farmers will refer to a person actively engaged in economic or livelihood activity of growing crops.

JOBS-TO-BE-DONE / PROBLEMS:

- The main problem is cultivating the food products for the growing population, providing a livelihood for farmers and protecting the environment.
- The potential impacts of heavy precipitation include crop damage, soil erosion, and an increase in flood which in turn can lead to injuries, drownings, and other flooding related effects on health.

TRIGGERS:

- Maturity levels of seeds and crop cultivation details are briefly explained.

EMOTIONS:

- BEFORE: Paying all debts Less crop production
- AFTER: Increase in sowing rate Economic growth

AVAILABLE SOLUTIONS:

- Aims in reducing the risks that raises in cultivating of crops.
- Developing agriculture, infrastructure including agricultural markets, cold storage for protecting of harvested crops, and warehouse.

CUSTOMER CONSTRAINTS:

- Prediction is never 100% and it is almost impossible to predict the future with certainty
- To determine how long it will be raining by the predicted rainfall amount.

BEHAVIOUR:

- Improper rainfall data may lead to loss of crops and farmlands.
- Inefficient rainfall data prediction could cause customers dissatisfaction.
- The atmosphere is constantly changing so their estimates becomes less reliable.

CHANNELS OF BEHAVIOUR:

- ONLINE: Gets notified about tomorrow's weather.
- OFFLINE: Agricultural products can be protected earlier.

PROBLEM ROOT CAUSE:

- Agriculture in India is dependent on the monsoon season because farmers are not well equipped with modern methods of manual irrigations.
- Too much of rain and floods damage crops in the fields and may drown livestock, destroy farmlands and damage stored agricultural products.

YOUR SOLUTION:

- A good balance of rain and proper irrigation can lead to faster-growing plants, which can cut down the length between seeding and harvest.
- It helps in making of decisions regarding the cropping pattern and sowing rate. • It helps in sustainable use of ground water resources.

Define CS, fit into CC

### 1.CUSTOMER SEGMENT(S)

- Farmers are the customers who cultivate
  different types of crops depending on the rainfall
- Farmers will refer to a person actively engaged in economic or livelihood activity of growing crops.

### 6.CUSTOMER CONSTRAINTS

- Prediction is never 100% and it is almost impossible to predict the future with certainty.
- To determine how long it will be raining by the predicted rainfall amount.

### 5.AVAILABLE SOLUTIONS

- Aims in reducing the risks that raises in cultivating of crops.
- Developing agriculture, infrastructure including agricultural markets, cold storage for protecting of harvested crops, and warehouse.

Explore AS, differentiate

Focus on J&P, tap into BE, understand RC

### 2.JOBS-TO-BE-DONE / PROBLEMS

- The main problem is cultivating the food products for the growing population, providing a livelihood for farmers and protecting the environment.
- The potential impacts of heavy precipitation include crop damage, soil erosion, and an increase in flood which in turn can lead to injuries, drownings, and other flooding related effects on health.

### 9.PROBLEM ROOT CAUSE

- Agriculture in India is dependent on the monsoon season because farmers are not well equipped with modern methods of manual irrigations.
- Too much of rain and floods damage crops in the fields and may drown livestock, destroy farmlands and damage stored agricultural products.

### 7.BEHAVIOUR

- Improper rainfall data may lead to loss of crops and farmlands.
- Inefficient rainfall data prediction could cause customers dissatisfaction.
- The atmosphere is constantly changing so their estimates becomes less reliable.

Focus on J&P, tap into BE, understand RC

### 3.TRIGGERS

- Maturity levels of seeds and crop cultivation details are briefly explained.

### 4.EMOTIONS:

**BEFORE:** Paying all debts
Less crop production
**AFTER:** Increase in sowing rate
Economic growth

### 10.YOUR SOLUTION

- A good balance of rain and proper irrigation can lead to faster-growing plants, which can cut down the length between seeding and harvest.
- It helps in making of decisions regarding the cropping pattern and sowing rate.
- It helps in sustainable use of ground water resources.

### 8.CHANNELS OF BEHAVIOUR

**ONLINE:**
Gets notified about tomorrow's weather.

**OFFLINE:**
Agricultural products can be protected earlier.

Extract online & offline CH of BE

# 4.REQUIREMENTS ANALYSIS

Requirements analysis or requirements engineering is a process used to determine the needs and expectations of a new product. It involves frequent communication with the stakeholders and end-users of the product to define expectations, resolve conflicts, and document all the key requirements.

One of the greatest challenges faced by any organization is to share the vision of the final product  the customers. Hence, a business requirements analysis involves a team effort of allthe key stakeholders, software developers, end-users, and customermanagers to achieve a shared understanding of what the product should do.

## 4.1.FUNCTIONAL REQUIREMENTS:

- Functional requirements are product features or functions that developers must implement to enable users to accomplish their tasks. So, it's important to make them clear both for the development team and the stakeholders. Generally, functional requirements describe system behavior under specific conditions.

Registration Process:

- Registration through Phone Number

Confirmation:

- Confirmation via OTP message

Updating Profile:

- Enter the personal details

Home Page:

Able to view the,

- Profile
- Crop details
- Rainfall prediction

Rainfall Prediction:

- Enter the month
- Enter the Year
- Click on predict

ML Model:

- The user data is sent to the Machine learning model.

Preprocessing data:

- Data exploration
- Feature selection
- Missing values
- Feature scaling
- Splitting of train and test data

Building ML Model:

- Random forest algorithm is applied
- Train the model using training data
- The model is evaluated with the test data

Result:

- Shows the predicted rainfall data.

### 4.2.NON FUNCTIONAL REQUIREMENTS

Nonfunctional Requirements (NFRs) define system attributes such as security, reliability, performance, maintainability, scalability, and usability. They serve as constraints or restrictions on the design of the system across the different backlogs.

Usability:

- It's a user-friendly application which enable people to use without any technical knowledge.

Security:

- User data will be protected from unauthorised access and the data are secured

Reliability:

- The application will operate effectively without causing any failure and errors, so maintance won't be big problem.

Performance:

- Overall performance of system is efficient to predict the rainfall with much speed without delay.

Availability :

- The availability of the application is that it will be active and available to all the users.

Scalability:

- The scalability of our system is one that can handle rapid changes to workloads.

# 5.PROJECT DESIGN

Designing a machine learning system is an iterative process. There are generally four main components of the process: project setup, data pipeline, modeling (selecting, training, and debugging your model), and serving.

## 5.1. DATA FLOW DIAGRAMS

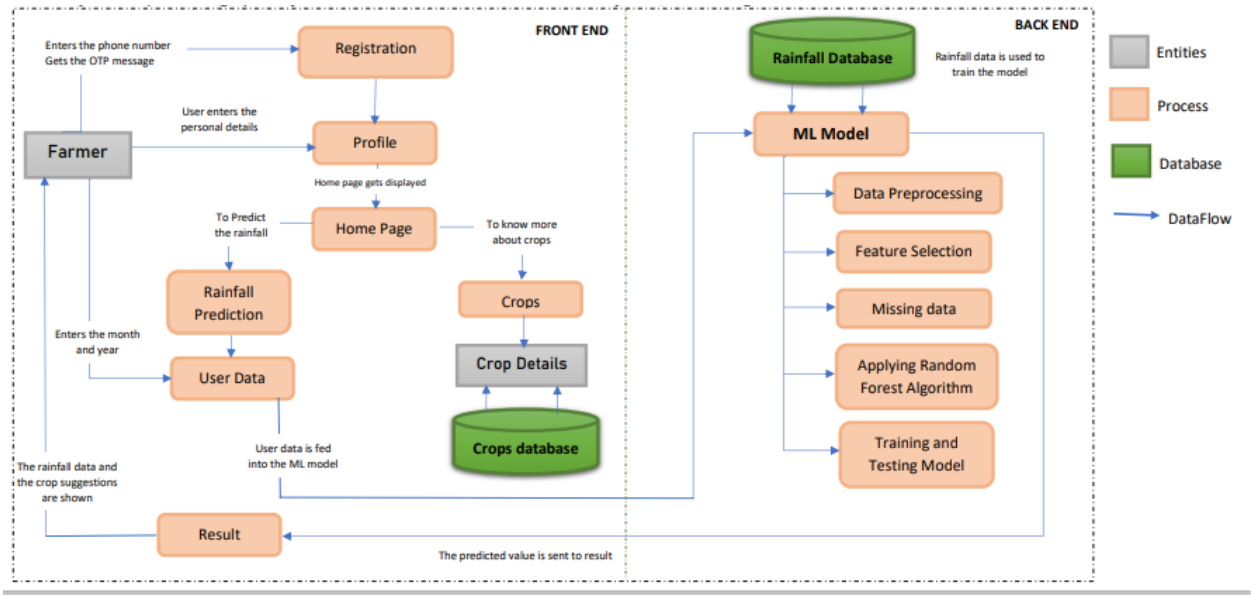
A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

- Users can able to register their profile by using their phone number by using their number thet can get a otp.
- After entering otp the user enters their personal details.
- Users can view their home page to know more about crops details and crop database.
- To predict the rainfall they enter month and the year to view rainfall data.
- Machine learning model predict the value is to senta result.
- The rainfall data and the crop suggestions are shown.

MACHINE LEARNING MODEL:

- Machine learning model used database to train the rainfall data,

☆ Data Preprocessing
☆ Feature Selection
☆ Mising Data
☆ Applying Snap Random Forest Algorithm




## 5.2.SOLUTION AND TECHNICAL ARCHITECTURE

Machine learning solutions are used to solve a wide variety of problems, but in nearly all cases the core components are the same. Whether you simply want to understand the skeleton of machine learning solutions better or are embarking on building your own, understanding these components - and how they interact - can help.

ML Solutions/Data Architects (meaning people responsible for the enterprise level infrastructure strategy of the ML/Data Science needs for the business) are primarily engineers with a lot of experience in building/supervising/managing complex infrastructure projects.

The technical architecture defines the various layers involved in the machine learning cycle and involves the major steps being carried out in the transformation of raw data into training data sets capable for enabling the decision making of a system.

Technical architecture includes the major components of the system, their relationships, and the contracts that define the interactions between the components. The goal of technical architects is to achieve all the business needs with an application that is optimized for both performance and security.

## 5.3.USER STORIES

- As a user, I can register for the application by using phone number
- As a user, I will receive OTP message, Once I have registered for the application
- As a user, I have to enter my personal details

- As a user, I can either click on rainfall prediction or crops button
- As a user, I can click on the “more action” button
- As a user, I have to enter the desired month and the year
- As a user, I can view the details of the crop

| User Type | Functional Requirement (Epic) | User Story Number | User Story / Task | Acceptance criteria | Priority | Release |
|---|---|---|---|---|---|---|
| Farmers (Mobile user) | Registration | USN-1 | As a user, I can register for the application by using phone number | I can access the profile and home page | High | Sprint-1 |
| | | USN-2 | As a user, I will receive OTP message, Once I have registered for the application | I can receive confirmation message | High | Sprint-1 |
| | Profile | USN-3 | As a user, I have to enter my personal details | I can access my home page | High | Sprint-1 |
| | Home page | USN-4 | As a user, I can either click on rainfall prediction or crops button | I can go to the desired page | Medium | Sprint-2 |
| | | USN-5 | As a user, I can click on the “more action” button | I can view my personal details | Medium | Sprint-2 |
| | Rainfall Prediction | USN-6 | As a user, I have to enter the desired month and the year | To know the rainfall on the given month | High | Sprint-3 |
| | Crops | USN-7 | As a user, I can view the details of the crops | To know more about the crop cultivations | Medium | Sprint-3 |

# 6.PROJECT PLANNING AND SCHEDULING

The process of planning primarily deals with selecting the appropriate policies and procedures in order to achieve the objectives of the project. Scheduling converts the project action plans for scope, time cost and quality into an operating timetable.

'Project Planning and Scheduling', though separate, are two sides of the same coin. In project management Fundamentally, 'Project planning' is all about choosing and designing effective policies and methodologies to attain project objectives. While 'Project scheduling' is a procedure of assigning tasks to get them completed by allocating appropriate resources within an estimated budget and time-frame.

**SPRINT PLANNING AND 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 using phone number | 3 | High | Kruthika |
| Sprint-1 | | USN-2 | As a user, I will receive OTP message, Once I have registered for the application | 4 | High | Pooja |
| Sprint-1 | Profile | USN-3 | As a user, I have to enter my personal Details. | 3 | High | Pooja |
| Sprint-2 | Home Page | USN-4 | As a user, I can either click on rainfall prediction or crops button | 6 | Medium | Suma |
| Sprint-2 | | USN-5 | As a user, I can click on the "more action" button | 4 | Medium | Sangeetha |
| Sprint-3 | Rainfall Prediction | USN-6 | As a user, I have to enter the desired month and the year | 6 | High | Sangeetha |
| Sprint-3 | Crops | USN-7 | As a user, I can view the details of the Crops. | 4 | Medium | Suma |
| Sprint-4 | Application development | USN-8 | As a user, I can use this application for predicting the rainfall to cultivate crops | 10 | High | Kruthika |

- A sprint estimation shows how much effort a series of tasks require. It's based on assumptions, requirements, and dependencies of a project.
- During sprint planning, we break the stories down into tasks, estimate those tasks, and compare the task estimates against our capacity. It's that, not points, that keep us from overcommitting in this sprint.
- Sprint planning is an event in the Scrum framework where the team determines the product backlog items they will work on during that sprint and discusses their initial plan for completing those product backlog items.

## 6.2.SPRINT DELIVERY SCHEDULE:

- Since sprints take place over a fixed period of time, it's critical to avoid wasting time during planning and development. And this is precisely where sprint scheduling enters the equation
- A sprint schedule is a document that outlines sprint planning from end to end. It's one of the first steps in the machine learning sprint planning process—and something that requires adequate research, planning, and communication

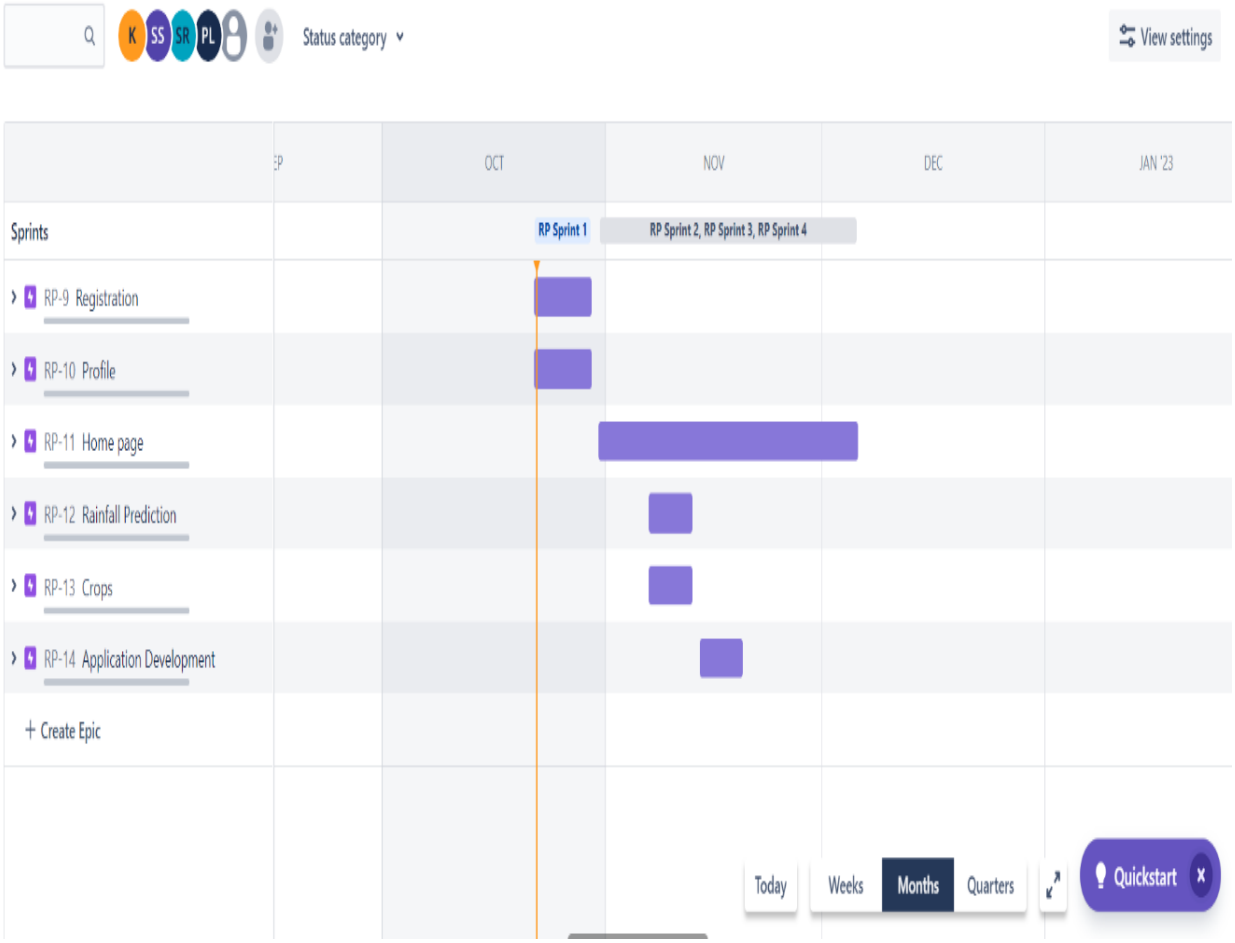
| Sprint | Total Story Points | Duration | Sprint Start Date | Sprint End Date (Planned) | Story Points Completed (as on Planned End Date) | Sprint Release Date (Actual) |
|---|---|---|---|---|---|---|
| Sprint-1 | 10 | 6 Days | 24 Oct 2022 | 29 Oct 2022 | 10 | 29 Oct 2022 |
| Sprint-2 | 10 | 6 Days | 31 Oct 2022 | 05 Nov 2022 | 10 | 05 Nov 2022 |
| Sprint-3 | 10 | 6 Days | 07 Nov 2022 | 12 Nov 2022 | 10 | 12 Nov 2022 |
| Sprint-4 | 10 | 6 Days | 14 Nov 2022 | 19 Nov 2022 | 10 | 19 Nov 2022 |

## 6.3.REPORTS FROM JIRA

- Reporting helps you track and analyze your team's work throughout a project.
- Jira Software has a range of reports that you can use to show information about your project, versions, epics, sprints, and issues.

STEPS:

1. Navigate to the project you want to report on.
2. From the project sidebar, select reports. The reports overview page displays.
3. Select a report from the overview or from the project sidebar to begin generating the report. View the list of reports below for more details of each report.


Status category
View settings
OCT
NOV
DEC
JAN '23
Sprints
RP Sprint 1
RP Sprint 2, RP Sprint 3, RP Sprint 4
RP-9 Registration
RP-10 Profile
RP-11 Home page
RP-12 Rainfall Prediction
RP-13 Crops
RP-14 Application Development
Create Epic
Today
Weeks
Months
Quarters
Quickstart

# 7. CODING AND SOLUTIONING

## 7.1TECHNOLOGY USED

## FEATURE 1

## Python:

Python is a interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

Since there is no compilation step, the edit-test-debug cycle is incredibly fast. Debugging Python programs is easy: a bug or bad input will never cause a segmentation fault. Instead, when the interpreter discovers an error, it raises an exception. When the program doesn't catch the exception, the interpreter prints a stack trace. A source level debugger allows inspection of local and global variables, evaluation of arbitrary expressions, setting breakpoints, stepping through the code a line at a time, and so on. The debugger is written in Python itself, testifying to Python's introspective power. On the other hand, often the quickest way to debug a program is to add a few print statements to the source: the fast edit-test-debug cycle makes this simple approach very effective.

**FLASK:**

Flask is a web application framework written in Python. It was developed by Armin Ronacher, who led a team of international Python enthusiasts called Poocco. Flask is based on the Werkzeg WSGI toolkit and the Jinja2 template engine.Both are

Pocco projects.

Flask is a web framework, it's a Python module that lets you develop web applications easily. It's has a small and easy-to-extend core: it's a microframework that doesn't include an ORM (Object Relational Manager) or such features.It does have many cool features like url routing, template engine. It is a WSGI web app framework.

**HTML:**

HTML stands for HyperText Markup Language. It is a standard markup language for web page creation. It allows the creation and structure of sections, paragraphs, and links using HTML elements such as tags and attributes. Developers use HTML code to design how a browser displays web page elements, such as text, hyperlinks, and media files. Internet navigation. Users can easily navigate and insert links between related pages and websites as HTML is heavily used to embed hyperlinks.HTML makes it possible to organize and format documents, similarly to Word.It's also worth noting that HTML is not considered a programming language as it can't create dynamic functionality. It is now considered an official web standard. The World Wide Web Consortium (W3C) maintains and develops HTML specifications, along with providing regular updates.

**CSS:**

Cascading Style Sheets (CSS) is a stylesheet language used to describe the presentation of a document written in HTML or XML (including XML dialects such as SVG, MathML or XHTML). CSS describes how elements should be rendered on screen, on paper, in speech, or on other media.Previously, the development of various parts of CSS specification was done synchronously, which allowed the versioning of the latest recommendations. You might have heard about CSS1, CSS2.1, or even CSS3. There will never be a CSS3 or a CSS4; rather, everything is now CSS without a version number.

After CSS 2.1, the scope of the specification increased significantly and the progress on different CSS modules started to differ so much, that it became more effective to develop and release recommendations separately per module. Instead of versioning the CSS specification, W3C now periodically takes a snapshot of the latest

stable state of the CSS specification and individual modules progress.JavaScript is a high-level, often just-in-time compiled language that conforms to the ECMAScript standard.[10] It has dynamic typing, prototype-based object-orientation, and first-class functions.It is multi-paradigm, supporting event-driven, functional, and imperative programming styles. It has application programming interfaces (APIs) for working with text, dates, regular expressions, standard data structures, and the Document Object Model (DOM).

## JAVA SCRIPT:

JavaScript is a high-level, often just-in-time compiled language that conforms to the ECMAScript standard. It has dynamic typing, prototype-based object-orientation, and first-class functions. It is multi-paradigm, supporting event-driven, functional, and imperative programming styles. It has application programming interfaces (APIs) for working with text, dates, regular expressions, standard data structures, and the Document Object Model (DOM).

JavaScript engines were originally used only in web browsers, but are now core components of some servers and a variety of applications. The most popular runtime system for this usage is Node.js.Although Java and JavaScript are similar in name, syntax, and respective standard libraries, the two languages are distinct and differ greatly in design.

### MYSQL:

MySQL is a relational database management system (RDBMS) developed by Oracle that is based on structured query language (SQL)A database is a structured collection of data. It may be anything from a simple shopping list to a picture gallery or a place to hold the vast amounts of information in a corporate network. In particular, a relational database is a digital store collecting data and organizing it according to the relational model. In this model, tables consist of rows and columns, and relationships between data elements all follow a strict logical structure. An RDBMS is simply the set of software tools used to actually implement, manage, and query such a database.

MySQL is integral to many of the most popular software stacks for building and maintaining everything from customer-facing web applications to powerful,

data-driven B2B services. Its open-source nature, stability, and rich feature set, paired with ongoing development and support from Oracle, have meant that internet-critical organizations such as Facebook, Flickr, Twitter, Wikipedia, and YouTube all employ MySQL backends.

**IBM:**

IBM, in full International Business Machines Corporation, leading American computer manufacturer, with a major share of the market both in the United States and abroad.

**IBM CLOUD STORAGE:**

IBM Cloud Object Storage provides a software-defined hyperscale storage solution which runs on premises. This industry-leading storage solution integrates with data on the edge, in your data center or on a private cloud.

**IBM WATSON STUDIO:**

IBM Watson Studio empowers data scientists, developers and analysts to build, run and manage AI models, and optimize decisions anywhere on IBM Cloud for Data. Unite teams, automate AI lifecycles and speed time to value on an open multicloud architecture.

**IBM MACHINE LEARNING:**

These insights subsequently drive decision making within applications and businesses, ideally impacting key growth metrics. As big data continues to expand and grow, the market demand for data scientists will increase, requiring them to assist in the identification of the most relevant business questions and subsequently the data to answer them.

**FEATURE 2**

**SNAP RANDOM FOREST CLASSIFIER:**

This class implements a random forest classifier using the IBM Snap ML library. It can be used for binary classification problems.

IBM has developed an efficient, scalable machine learning library that enables very fast training of generalized linear models. Using this library, clients can remove training time as the bottleneck for machine learning workloads, paving the way to a range of new applications.

Snap ML is offered through PowerAI (current version 1.6.0, released March 22nd), where it exposes two APIs, a Python one for single-machine use or distributed use across a cluster, and a Spark one for use in an Apache Spark cluster.

**MAIN FUNCTION CODING:**

```
from flask import Flask, render_template, request, redirect, url_for, session
from flask_mysqldb import MySQL
import MySQLdb.cursors
import re
import os
from twilio.rest import Client
from werkzeug.exceptions import HTTPException
import requests
# NOTE: you must manually set API_KEY below using information retrieved from your
IBM Cloud account.
API_KEY = "4pewzbYZOBMZuMhd1PDoWyHOx4J4oUGn6eZPiycEvdJJ"
token_response = requests.post('https://iam.cloud.ibm.com/identity/token',
data={"apikey":
 API_KEY, "grant_type": 'urn:ibm:params:oauth:grant-type:apikey'})
mltoken = token_response.json()["access_token"]



header = {'Content-Type': 'application/json', 'Authorization': 'Bearer ' + mltoken}

app = Flask(_name_)
```

```
# Change this to your secret key (can be anything, it's for extra protection)
app.secret_key = 'your secret key'
# Enter your database connection details below
app.config['MYSQL_HOST'] = 'localhost'
app.config['MYSQL_USER'] = 'root'
app.config['MYSQL_PASSWORD'] = '5005kruthi'
app.config['MYSQL_DB'] = 'pythonlogin'


# Intialize MySQL
mysql = MySQL(app)


@app.route('/', methods=['GET', 'POST'])
def login():
    # Output message if something goes wrong...
    msg = ''
    # Check if "username" and "password" POST requests exist (user submitted form)
    if request.method == 'POST' and 'log_mail' in request.form and 'pswd' in request.form:
        # Create variables for easy access
        log_mail = request.form['log_mail']
        pswd = request.form['pswd']
        # Check if account exists using MySQL
        cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
        cursor.execute('SELECT * FROM register WHERE email = %s AND pswd = %s',
(log_mail, pswd,))
        # Fetch one record and return result
        profile = cursor.fetchone()
        # If account exists in accounts table in out database
        if profile:
            # Create session data, we can access this data in other routes
```

```
            session['loggedin'] = True
            session['id'] = profile['id']
            session['log_mail'] = profile['email']
            # Redirect to home page
            return redirect(url_for('home'))
        else:
            # Account doesnt exist or username/password incorrect
            msg = 'Incorrect email/password!'
    # Show the login form with message (if any)
    return render_template('login.html', msg=msg)


@app.route('/logout')
def logout():
    # Remove session data, this will log the user out
   session.pop('loggedin', None)
   session.pop('id', None)
   session.pop('log_mail', None)
   # Redirect to login page
   return redirect(url_for('login'))

@app.route('/getotp', methods=['GET', 'POST'])
def getotp():
    msg=''
    if request.method == 'POST' and 'phone' in request.form and 'uname' in request.form :
      # Create variables for easy access
      global phone,uname
      phone = request.form['phone']
```

```
    uname = request.form['uname']
    cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
    cursor.execute('SELECT * FROM registration WHERE phone= %s', (phone,))
    account = cursor.fetchone()
    if account:
        msg = 'Account already exists!'
    else:
        #cursor.execute('INSERT INTO registration VALUES (NULL, %s)', (phone,))
        #mysql.connection.commit()
        try:
            account_sid ='ACf605690bd4d15edaf43560e3be75d622'
            auth_token = '050436e3811cf120fa1b12f77fd0fae2'
            client = Client(account_sid, auth_token)
            verification = client.verify \
                        .v2 \
                        .services('VAf924698b5a809f7041269be39e245851') \
                        .verifications \
                        .create(to=phone, channel='sms')
            if verification.status=='pending':
                msg='OTP Sucessfully sent!'
            else:
                msg='Check Your Phone Number'
        except Exception:
            msg='Enter Your PhoneNumber Correctly'
    return render_template('signup.html', msg=msg)
```

```
@app.route('/register', methods=['GET', 'POST'])
def register():
    # Output message if something goes wrong...
    msg = ''
    # Check if "username", "password" and "email" POST requests exist (user submitted form)
    if request.method == 'POST' and 'otp' in request.form:
        # Create variables for easy access
        otp= request.form['otp']
        account_sid ='ACf605690bd4d15edaf43560e3be75d622'
        auth_token = '050436e3811cf120fa1b12f77fd0fae2'
        client = Client(account_sid, auth_token)

        verification_check = client.verify \
                             .v2 \
                             .services('VAf924698b5a809f7041269be39e245851') \
                             .verification_checks \
                             .create(to=phone, code=otp)

        if verification_check.status=='approved':
            cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
            cursor.execute('INSERT INTO registration VALUES (NULL, %s, %s)', (phone,uname,))
            mysql.connection.commit()
            return redirect(url_for('profile'))
        else:
            msg = 'Incorrect OTP'
```

```
    # Show registration form with message (if any)
    return render_template('signup.html', msg=msg)


@app.route('/profile', methods=['GET', 'POST'])
def profile():
    # Output message if something goes wrong...
    msg = ''
    # Check if "username", "password" and "email" POST requests exist (user submitted form)
    if request.method == 'POST' and 'username' in request.form and 'pswd' in request.form and 'email' in request.form and 'location' in request.form and 'occupation' in request.form and 'con_pswd' in request.form:
        # Create variables for easy access
        username = request.form['username']
        email = request.form['email']
        location = request.form['location']
        occupation= request.form['occupation']
        con_pswd = request.form['con_pswd']
        pswd = request.form['pswd']

        # Check if account exists using MySQL
        cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
        cursor.execute('SELECT * FROM register WHERE email = %s', (email,))
        register = cursor.fetchone()
        # If account exists show error and validation checks
```

```
        if register:
            msg = 'Account already exists!'
        elif not re.match(r'[^@]+@[^@]+\.[^@]+', email):
            msg = 'Invalid email address!'
        elif pswd!=con_pswd:
            msg='Password did not Match'
        elif not username or not pswd or not email:
            msg = 'Please fill out the form!'
        else:
            # Account doesnt exists and the form data is valid, now insert new account into
accounts table
            cursor.execute('INSERT INTO register VALUES (NULL, %s, %s, %s,%s,%s,%s,%s)',
(username,email,location,occupation,pswd,con_pswd,phone,))
            mysql.connection.commit()
            msg = 'You have successfully registered!'
            cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
            cursor.execute('SELECT * FROM register WHERE email = %s AND pswd = %s',
(email, pswd,))
            # Fetch one record and return result
            profile = cursor.fetchone()
            # If account exists in accounts table in out database
            if profile:
                # Create session data, we can access this data in other routes
                session['loggedin'] = True
                session['id'] = profile['id']
                session['log_mail'] = profile['email']
                # Redirect to home page
```

```
            return redirect(url_for('home'))

    elif request.method == 'POST':
        # Form is empty... (no POST data)
        msg = 'Please fill out the form!'
    # Show registration form with message (if any)
    return render_template('profile-page.html', msg=msg)

@app.route('/home', methods=['GET', 'POST'])
def home():
    if 'loggedin' in session:
        # User is loggedin show them the home page
        return render_template('popup.html')
    # User is not loggedin redirect to login page
    return redirect(url_for('login'))

@app.route('/review', methods=['GET', 'POST'])
def review():
    # Check if "username", "password" and "email" POST requests exist (user submitted
form)
    if request.method == 'POST' and 'review' in request.form:
        # Create variables for easy access
        review = request.form['review']
        cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
        cursor.execute('INSERT INTO feedback VALUES (NULL, %s)', (review,))
        mysql.connection.commit()
        msg="Thank You for the valuable feedback!"
```

```
        return render_template('popup.html', msg=msg)
    else:
        msg='Fill the form!'
        return render_template('popup.html', msg=msg)
@app.route('/rate', methods=['GET', 'POST'])
def rate():
    if request.method == 'POST' and 'rate' in request.form:
        # Create variables for easy access
        rate = request.form['rate']
        cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
        cursor.execute('INSERT INTO rate VALUES (NULL, %s)', (rate,))
        mysql.connection.commit()
        msg="Thanks for Rating Us!"
        return render_template('popup.html', msg=msg)
    else:
        msg='Fill the form'
        return render_template('popup.html', msg=msg)
@app.route('/resetotp', methods=['GET', 'POST'])
def resetotp():
    msg=''
    if request.method == 'POST' and 'mobile' in request.form:
      # Create variables for easy access
      global mobile
      mobile = request.form['mobile']
      cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
      cursor.execute('SELECT * FROM register WHERE phone= %s', (mobile,))
      account = cursor.fetchone()
```

```
    if account:
        try:
            account_sid ='AC5a17fbe4f1136a9420b131184eccf02f'
            auth_token = '8976233f3e08dfdf8fc99946ceb3858c'
            client = Client(account_sid, auth_token)
            verification = client.verify \
                        .v2 \
                        .services('VA258829db0002722215f1024b3e94c64b') \
                        .verifications \
                        .create(to=mobile, channel='sms')
            if verification.status=='pending':
                msg='OTP Sucessfully sent!'
            else:
                msg='Check Your Phone Number'
        except Exception:
            msg='Enter Your PhoneNumber Correctly'
    else:
        msg='You dont Have an Account'
    return render_template('forgot.html', msg=msg)

@app.route('/otpverify', methods=['GET', 'POST'])
def otpverify():
    # Output message if something goes wrong...
    msg = ''
    # Check if "username", "password" and "email" POST requests exist (user submitted
form)
    if request.method == 'POST' and 'otp' in request.form:
```

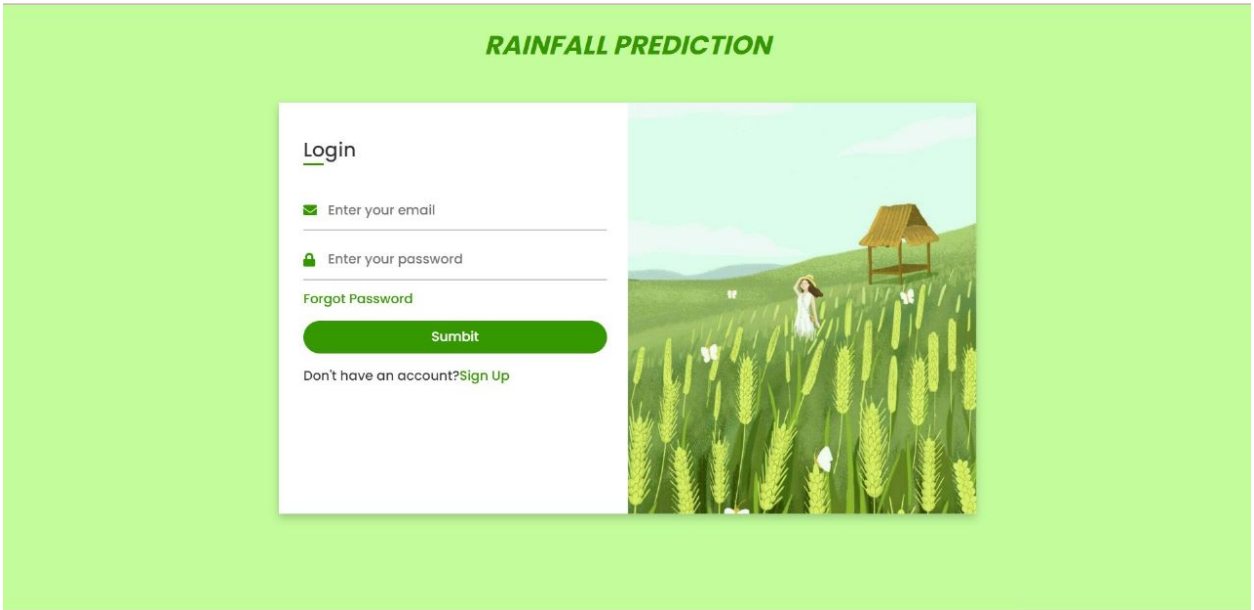
```
        # Create variables for easy access
        otp= request.form['otp']
        account_sid ='AC5a17fbe4f1136a9420b131184eccf02f'
        auth_token = '8976233f3e08dfdf8fc99946ceb3858c'
        client = Client(account_sid, auth_token)

        verification_check = client.verify \
                            .v2 \
                            .services('VA258829db0002722215f1024b3e94c64b') \
                            .verification_checks \
                            .create(to=mobile, code=otp)

        if verification_check.status=='approved':
            return redirect(url_for('setpswd'))
        else:
            msg = 'Incorrect OTP'

    # Show registration form with message (if any)
    return render_template('forgot.html', msg=msg)

@app.route('/setpswd', methods=['GET', 'POST'])
def setpswd():
    msg=''
    if request.method == 'POST' and 'pswd' in request.form and 'con_pswd' in
request.form:
        pswd= request.form['pswd']
        con_pswd= request.form['con_pswd']
```

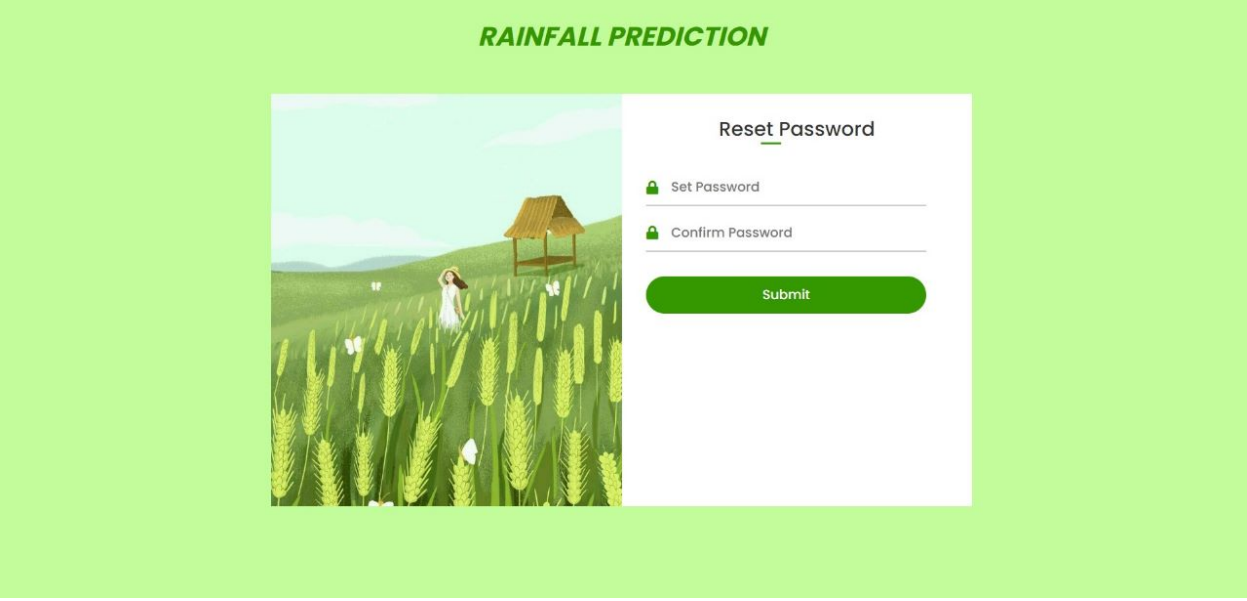
```
        if pswd==con_pswd:
            cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
            cursor.execute('UPDATE register SET pswd=%s WHERE phone= %s', (pswd,mobile
,))
            cursor.execute('UPDATE register SET con_pswd=%s WHERE phone= %s',
(con_pswd,mobile,))
            mysql.connection.commit()
            return redirect(url_for('login'))
        else:
            msg='Password did not Match'

    return render_template('reset.html', msg=msg)

@app.route('/crop', methods=['GET', 'POST'])
def crop():
    return render_template('crop.html')

@app.route('/rainfall', methods =['POST','GET'])
def rainfall():
    msg=''
    list=[]
    if request.method == 'POST' and 'month' in request.form and 'location' in request.form
and 'year' in request.form:
        month =request.form["month"]
        location= request.form["location"]
        year= request.form["year"]
        if len(str(year))==4:
```

```
        t=[[location,int(year),month]]
        payload_scoring = {"input_data":  [{"fields": [["l","y","m"]], "values":t}]}
        response_scoring = requests.post('https://us-
south.ml.cloud.ibm.com/ml/v4/deployments/c55037c7-0ca7-4482-bbe3-
d2c6aaec0811/predictions?version=2022-11-18', json=payload_scoring,
                                         headers={'Authorization': 'Bearer ' + mltoken})
        print("Scoring response")
        print(response_scoring.json())
        pred= response_scoring.json()
        output=pred['predictions'][0]['values'][0][0]
        print(output)
        res=float("{:.2f}".format(output))
        out=str(res)+"cm"
        if res>=0 and res<=15:
           list=["Sponge gourd","Bottle Gourd","Coriander","Daisy","Carrot"]
        elif res>=16 and res<=50:
           list=["Cotton", "Sugarcane","Barley","Cucumber", "Pumpkin", "Bitter gourd", "Chili",
"Beans", "Cauliflower", "Cluster beans", "Soyabean", "Green amarnath", "Kasturi methi",
"Tinda" ,"Custard apple", "Watermelon", "Rose", "Sunflower", "Lilly"]
        elif res>51 and res<=75:
           list=["Wheat", "Millet", "Ragi", "Maize", "Cotton", "Sugarcane", "Groundnuts",
"Barley", "Mustard", "Brinjal", "Onion" , "Capsicum", "Spring Onion", "Beetroot" , "Spinach",
"Fenugreek", "Garlic", "Soyabean", "Mushroom", "Broad beans", "Green amaranth", "Kasturi
methi", "Mango", "Custard apple", "Watermelon", "Sunflower", "Brama kamal", "Iris"]
        elif res>76 and res<=100:
           list=["Wheat", "Millets", "Ragi", "Maize", "Cotton", "Sugarcane", "Groundnuts",
"Barley", "Mustard", "Tomato", "Brinjal", "Onion", "Capsicum", "Ladies finger","Spring
```

```
Onion", "Beetroot", "Spinach", "Fenugreek","Garlic", "Soyabean", "Broccoli", "Broad beans",
"Green amarnath", "Mango", "Apple", "Pine apple", "Custard apple", "Lotus", "Jasmin"]
        elif res>101 and res<=150:
            list=["Ragi", "Sugarcane", "Capsicum", "Beetroot", "Spinach", "Ridge gourd",
"Chayote", "Pointed gourd","Drumstick", "Spine gourd", "Green amarnath", "Turnip",
"Mango", "Apple", "Pine apple", "Lotus"]
        elif res>151 and res<=200:
            list=["Rice", "Chayote", "Spine gourd", "Green amarnath", "Turnip", "Mango",
"Banana", "Pine apple", "Lotus", "Jute"]
        elif res>201 and res<=250:
            list=["Rice" , "Spine gourd", "Green amarnath", "Turnip", "Mango", "Banana", "Pine
apple"]
        elif res>251 and res<=300:
            list=["Rice", "Ginger","Green amarnath","Turnip","Mango","Banana","Orchid"]
        else:
            list=["Rice","Mango"]
        return render_template("pred.html",msg=out,list=list)
      else:
        msg="Enter Year correctly"
   return render_template('rainfall.html',msg=msg)

if _name_ == '_main_' :
   app.run(debug= False)
```

**SAMPLE OUTPUT:**

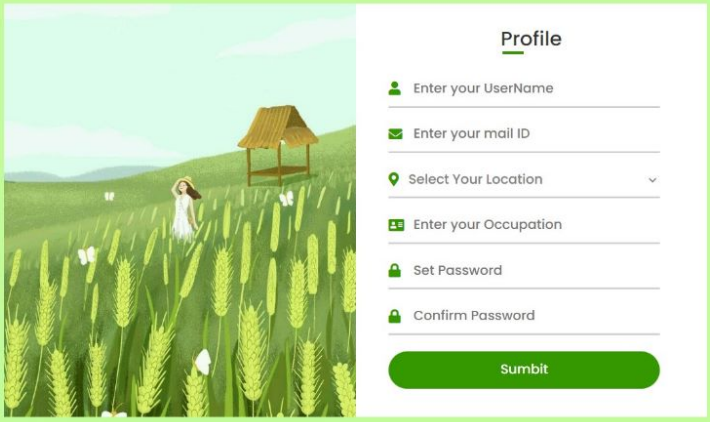


***RAINFALL PREDICTION***

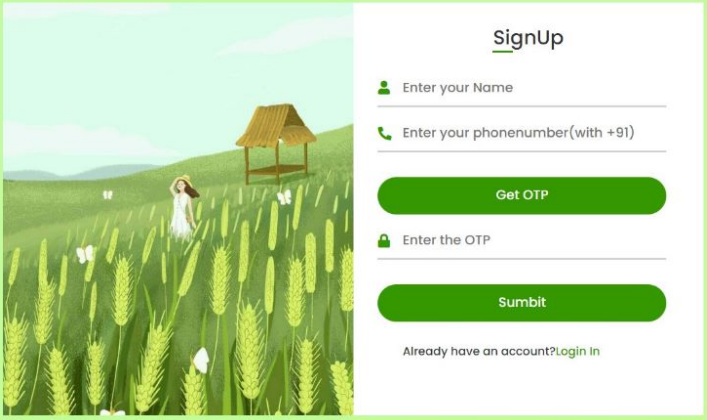

***RAINFALL PREDICTION***

***RAINFALL PREDICTION***

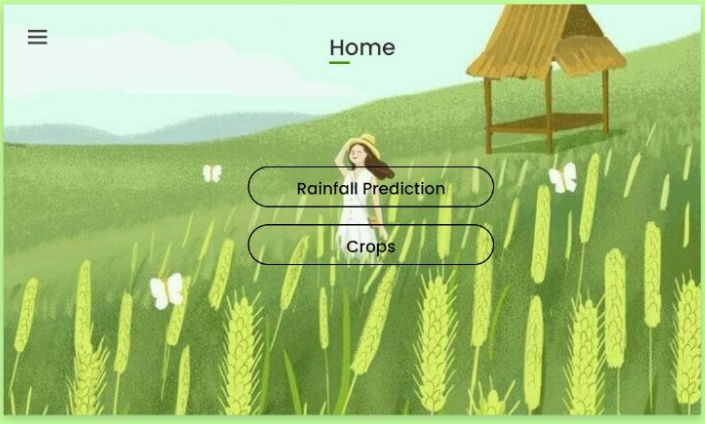


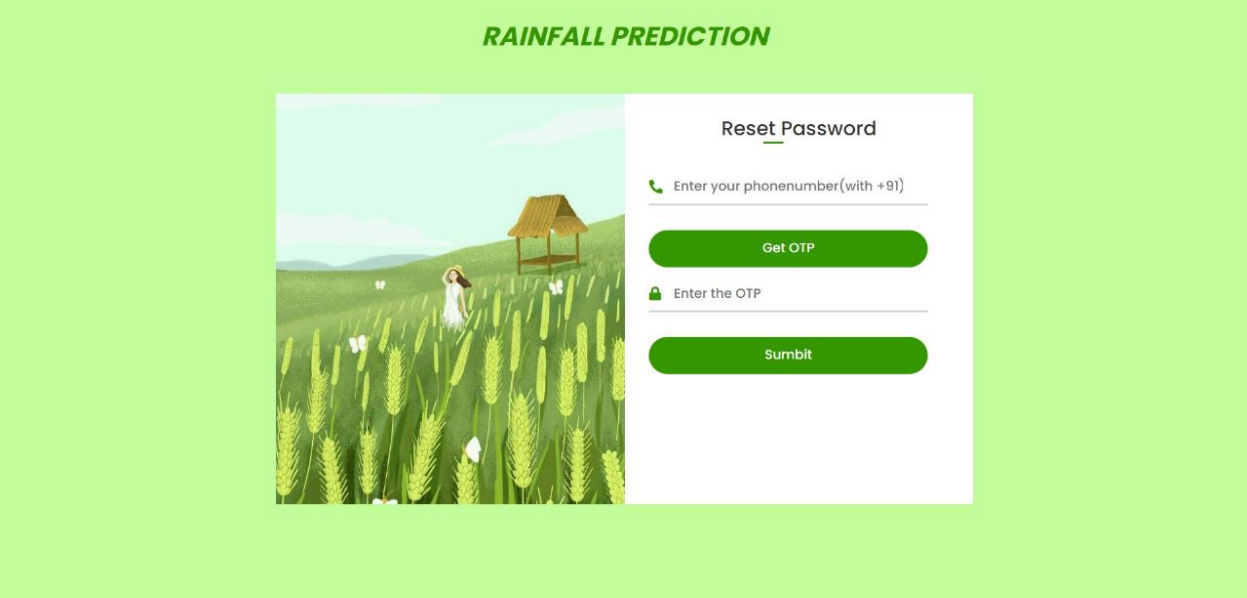

***RAINFALL PREDICTION***

**RAINFALL PREDICTION**




**RAINFALL PREDICTION**






**DATABASE SCHEMA:**

**REGISTRATION TABLE:**

```
CREATE TABLE `registration` (
 `id` int NOT NULL AUTO_INCREMENT,
 `phone` varchar(15) DEFAULT NULL,
 `name` varchar(50) DEFAULT NULL,
 PRIMARY KEY (`id`)
) ENGINE=InnoDB AUTO_INCREMENT=3 DEFAULT CHARSET=utf8mb3
```

**REGISTER TABLE:**

```
CREATE TABLE `register` (
  `id` int NOT NULL AUTO_INCREMENT,
  `username` varchar(50) DEFAULT NULL,
  `email` varchar(50) DEFAULT NULL,
  `location` varchar(50) DEFAULT NULL,
  `occupation` varchar(50) DEFAULT NULL,
```

```
    `pswd` varchar(50) DEFAULT NULL,
    `con_pswd` varchar(50) DEFAULT NULL,
    `phone` varchar(15) DEFAULT NULL,
  PRIMARY KEY (`id`)
) ENGINE=InnoDB AUTO_INCREMENT=3 DEFAULT CHARSET=utf8mb3
```

**FEED BACK:**

```
CREATE TABLE `feedback` (
  `id` int NOT NULL AUTO_INCREMENT,
  `review` varchar(300) DEFAULT NULL,
  PRIMARY KEY (`id`)
) ENGINE=InnoDB AUTO_INCREMENT=3 DEFAULT CHARSET=utf8mb3
```

**RATING TABLE:**

```
CREATE TABLE `rate` (
  `id` int NOT NULL AUTO_INCREMENT,
  `rating` int DEFAULT NULL,
  PRIMARY KEY (`id`)
) ENGINE=InnoDB AUTO_INCREMENT=2 DEFAULT CHARSET=utf8mb3
```

# 8. RESULTS

Analysis of monthly rainfall data from GMet between 1985 and 2014 established a pattern of variability. Three data sets, described as Climate Assessment Decade (CAD), of 10 years, ranging from 1985 to 1994, 1995 to 2004, and 2005 to 2014, were categorised to allow comparison of variation in rainfall distribution in the area. As the area experiences a bimodal rainfall regime, the CAD for each data set is further grouped into the major season, between March and July, and minor season, between September and November. This grouping was intended to examine the comparative basis for the extent of variability. The result shows a trend in seasonal (major and minor) variability in rainfall distribution.The major season for the first decade under consideration (1985-1994) recorded an average rainfall of 1449.94 mm compared to the third decade (1995-2004) which recorded a significantly lower rainfall (1278.58 mm) and 1098.94 mm between 2005 and 2014. In other words, there was a decrease in rainfall distribution in the major seasons over the 30-year period. Similarly, the minor season for each CAD recorded increased variability (Figures 2 and 3), 1374.98 mm, 967.50 mm, and 1117.82 mm for 1985-1994, 1995-2004, and 2005-2014, respectively, with the highest incidence of variability observed in the first and second decades of the period under consideration.

## 8.1.PERFORMANCE METRICS:

The development, grow and yield of crops is influenced by the seasonal patterns in rainfall and temperature; and therefore, any future alteration of these, may have significant impacts on agricultural production (Knox et al. 2010a; Falloon and Betts 2010; Murphy et al. 2009). Future projections of increased daily temperature and humidity in the atmosphere can also increase the risk of agricultural pests and diseases as well as deteriorate the land available for agricultural activities due to sea rise level.

## 9.ADVANTAGES

- Contributes substantially to vegetation, notably for agricultural production, as most farmers in third-world regions rely solely on rain
- High temperature and rainfall allow many different crops to grow.
- Rainfall is very important to farmers to cultivate all types of crops in seasonal process.
- Rainfall can also determine how fast a crop will grow from seed, including when it will be ready for harvesting.
- A good balance of rain and proper irrigation can lead to faster-growing plants, which can cut down on germination time and the length between seeding and harvest.
- The soil can also start to collect bacteria, mold, and fungus, which can then be absorbed by the plant.
- A good balance of rain and proper irrigation can lead to faster-growing plants, which can cut down on germination time and the length between seeding and harvest.
- There are 5 best instances, rice, maize, Basra, corn, & peas, and many others. We can say that rain is key to good agriculture.

## 9.1 DISADVANTAGES

- There is the potential of mild to moderate depression because cloudy and gloomy weather might dampen thoughts and feelings
- If the average rainfall is much lower or higher than the ideal, it can lead to significant problems, from drowned crops to lower yields.
- Soil is also greatly affected by rainfall. If it is too wet or too dry, nutrients in the soil can run off and not make it to the plants' roots, leading to poor growth and overall health.
- The crops are dependent on water during their entire lifecycle in order to survive and thrive.So without rainfall crops cant survive.
- Heavy rainfall may leads to loss of crops and it wont give profit to the farmers.
- Irregular rainfall can have many impacts like destruction of crops and farms, damage of property.
- The potential impacts of heavy precipitation include crop damage, soil erosion, and an increase in flood risk due to heavy rain

# 10.CONCLUSION

Agricultural economists and planners in India believed that the quantum of monsoon rainfall in a particular year (or its departure from mean values) would mainly determine the annual agricultural growth rate that the country would achieve in that year. This was basically the logical extension of the valid concept that agricultural output of the country in a particular year can change with the monsoon rainfall which ensures sufficient moisture for the kharif (monsoon) crops and adequate inflows in reservoirs and replenishment of aquifers for irrigation. The absence of longterm data on average annual rainfall on the country's landmass, however, made it difficult to test this hypothesis. But such views ignore the fact that at times there can be wide fluctuations in agricultural outputs between two consecutive years due to large variations in the corresponding rainfall, and as a result, the value of agricultural outputs can be less than that of the previous year. Poor agricultural growth performance observed in certain years was attributed to the inadequate rainfall of that year and high growth rates observed in certain other years was attributed to effective policy interventions and good monsoon. The influence of the preceding year's rainfall in determining agricultural growth performance of a year was largely ignored.Wide variation in agricultural growth rates between years which experienced more or less the same quantum of rainfall suggests that there are factors other than annual rainfall which influence the agricultural growth rates.

The variation depends on the timing resolution with which these water supply variables are considered by the modelers. The results show that the coarser the timing resolution, the more the impacts on agricultural income are underestimated. For example, as presented in the event of a 30% cut in rainfall, while the yearly model would predict a 10.2% drop in net-revenue, the seasonal model predicts a drop of 10.4% and the monthly model, 11.6%. And as the cuts get deeper, the differences in the predictions become higher. For example, in the event of a 50% cut in rainfall, the yearly, seasonal and monthly models would predict a drop of 20%, 28% and 34% respectively. Between the yearly and the monthly model, impacts are higher in absolute terms, in the latter by 14%.

These results have clear implications for the study of water scarcity impacts on agricultural income and ultimately on the design of cost-effective public policies that aim to lift up farmers from poverty conditions prevalent in many rural areas of the world that are subject to water scarcity. While the spatial extent of the current empirical example is too small to be scaled up, results do indicate that the timing with respect to

climatic variables such as rainfall and water supply, in this case water stored in reservoirs, may significantly influence the precise quantification of water scarcity impacts.

Rainfall Prediction is the application area of data science and machine learning to predict the state of the atmosphere. It is important to predict the rainfall intensity for effective use of water resources and crop production to reduce mortality due to flood and any disease caused by rain. This paper analyzed various machine learning algorithms for rainfall prediction.

## 11.FUTURE SCOPE

- We have used Indian dataset with states and its average rainfall..we can expand it to all districts seperately.
- It can be expanded to predict everyday rainfall.
- Our project is a web application,,it can be made into Android application.
- We can enable user to register from Google and from other social media.

## 12.APPENDIX

### PROJECT DEMO LINK:

https://drive.google.com/file/d/12LhM5qQ5ljxmSf14N4z4_SEG6Jt9fPeO/view?usp=drivesdk

**GITHUB LINK:**

https://github.com/IBM-EPBL/IBM-Project-24055-1659936361